3.1 **Introduction**

Research in common parlance, means a journey from the known to unknown. It has been described as an art of methodical examination of a specific problem, in order to establish the facts. The term ‘Research’ has been defined by many authors with various angles.

As stated by (Grinnell, 1993), “Research is a structured inquiry that utilizes acceptable scientific methodology to solve problems and create new knowledge which can be generally applied”. Research is usually categorized as per the application, objectives and enquiry mode of the research being conducted etc.

‘Research Methodology’ is a systematic way of solving a research problem. It not only describes various research methods, but also explains the logic behind using those methods. Statistics and methods of research help a business firm in gaining a competitive edge in the market over its rivals. It also in turn leads to reduction of the business risk, thereby making the risk controllable. (Burns & Burns, 2008).

This chapter has stated the way in which the research has been carried out by the researcher. It gives detailed information about data collection and the way in which it has been processed. Researcher has used secondary as also primary data for the present study.

3.2 **Research Philosophy for the present study**

This research was conducted by the researcher to study the indicators that prompt the need of forensic accounting for a sector and the companies in that sector. As the financial ratios have proved to be of great help in fraud detection so far; the researcher also diverted the efforts to find if financial ratios can be of value in prevention and early detection of fraud possibilities as envisaged in this particular research.

The researcher conducted the present study by using primary as well as secondary data.
Key features of the research undertaken by the researcher were as follows-

1) This was a correlational research. The researcher has studied the relationship between financial ratios and their ability to indicate fraud possibilities.

2) The study covered 17 prominent sectors of Indian economy and thus, gave a broader view of the overall economy.

3) It was a study of sector-specific red flags by using financial statements of selected public companies. As these statements are easily accessible by public at large, this study can be practiced very easily by its users in future.

4) It was a ‘quantitative’ study of identifying vulnerability of companies to fraud risk, hence was very objective in nature.

5) The researcher used a mix of primary and secondary data for the present study in order to cover various dimensions of the subject.

3.3 Research Gap

The extensive literature review conducted by the researcher brought forward a few glaring gaps in the existing literature in the field of fraud risk control and forensic accounting. These gaps created scope for further research in the area of forensic accounting. Researcher has closed some of the gaps so found.

Research in forensic accounting by big chartered accountancy firms like KPMG, PWC and Deliotte is significant in Indian context. Companies Act 2013, ‘Serious Fraud Investigations Office’ established by Ministry of Corporate Affairs, forensic wing in banks and seriousness of Reserve Bank in that process, training initiatives taken by ICAI and ACFE etc. have contributed to the wave of forensic accounting in India.

Few researchers like (Mehta et.al. 2012), (Sharma & Panigrahi, 2012) etc. have done extensive research on the literature on forensic accounting methods, models of fraud detection, scams that have taken place etc. However, realization of the need of forensic accounting and awareness about its utility in fraud control activity is yet limited in India.
This research decided to fill this gap by studying the indicators which would prompt the need of forensic accounting. This need identification was done by recognizing the vulnerability of companies to fraud risk.

Lot of research has been done on fraud detection. But fraud prevention initiatives have been relatively limited. Literature has underlined fraud prevention, transparent financial reporting, robust fraud risk management system etc. as the aspects critical to the survival and growth of any organization. Understanding the indicators that contribute to the risk of fraud is the first step of fraud prevention (Chersan, 2009). Thus, the researcher, in this study, has studied indicators that show vulnerability of a company towards fraud risk. Vulnerability to fraud risk creates a need to mitigate this risk. Risk mitigation is can be done very effectively through the use forensic accounting services.

Countries like Nigeria, USA, and Turkey etc. are much ahead of us in the context of research on this topic. They have studied fraud and non-fraudulent companies with the help of advanced statistical tools like cluster analysis, fuzzy logic, artificial neural network, Logit and Probit analysis etc. However, not much research has been found on sector specific fraud indicators or red flags, in quantitative terms. Available research on quantitative red flags has been done, again, by comparing fraud and non-fraud firms.

In India, labelling a company as fraud or non-fraud is not possible due to lack of precise information on this issue. As such, comparison based study was not be picked up by the researcher.

The researcher, in this study, has taken a wider cross section of Indian economy and has picked up 2595 companies across 17 significant sectors. Z scores have been used in the study to classify companies as red flag and green flag companies.

Study on qualitative red flags of fraud is a lot more as compared to quantitative study of identifying red flags of fraud. Two most relevant models giving quantitative indicators are Altman model (1968) and Beniesh (1999) model which have the ability to find out the likelihood of fraudulent financial reporting related to a company.
Others researchers like (Persons, 1995), (Green & Choi, 1991) also contributed in this area by use of financial ratios as detection tools or tools to detect factors associated with FFS. However, not much has been studied to prevent frauds. Again, all if it was based on fraud and non-fraud comparison; not feasible in Indian context. Researcher in the study under consideration, has studied the ‘quantitative indicators’ based on financial statements. It has been understood from the review of literature that financial statement frauds are most costly in terms of fraud losses. Researcher found these frauds as crucial to success and long term sustainability of companies. Thus the present study focused on financial statement frauds.

3.4 Research Problem

Literature review conducted shows a dire need to research further in the area of early detection of frauds. Specially talking about Indian scenario, it is seen that frauds are on the rise and yet not much is being done to strategically prevent them. Qualitative indicators are relatively subjective in nature. For example, an employee living beyond his means is an important qualitative indicator that speaks about possibility of that employee being a fraudster. Or Profit figures are too good to be true is another such qualitative indicator. But these do not quantify the fraud risk. Numbers enable a person understand the situation better than just words. Thus, quantitative indications, whether in financial terms or in statistical terms, should be used for this purpose.

It was found from the literature that red flags or fraud indicators do not mean that fraud necessarily exists in an organization. However they work as “Fraud Risk Identifiers” and speak about the susceptibility of an organization to frauds. These signals can guide managers on fraud risk mitigation before any fraud takes a massive form. They also would help management in installing sufficient and appropriate internal controls, both operational and financial.

A fraud survey report by ACFE in 2012 stated that each year a median loss of 5% of the annual revenue is faced by organizations across the globe including India. The problem of misstatement, whether due to frauds or errors, is huge with no formal solution.
Deloitte India Fraud Survey report of 2014 based on a sample of 400 executives mentioned that, only 38% respondents’ organize trainings of fraud risk assessment. This speaks about the laid back attitude towards fraud risk assessment in particular and fraud prevention in general. Report says that companies need to invest in fraud control mechanisms for long term sustainability.

Thus, identifying indicators of fraud risk that companies face and ensuring proactive approach of managers and directors towards fraud risk control was the prime research problem that researcher observed in the review of literature. It was seen that concept of red flags in forensic accounting shows a path towards fraud risk assessment and control. Forensic accounting in this sense is the way as well as the means of fraud risk mitigation.

3.5 Research Question
The researcher wanted to find whether published financial statements can give indications about the possibility of fraud, thereby, necessitating the use of forensic accounting principles and practices.

The focus of the study was on identification of indicators with the help of financial ratios and their analysis. Literature has shown that a lot of financial ratios have predictive ability as far as falsification of books of accounts and financial statement frauds are concerned. The researcher studied these ratios to examine if these ratios could be effective in Indian context.

Researcher was also interested in finding sector-specific indicators, which in turn would narrow down the task of fraud prevention. If the research can bring forth specific areas which are relatively more vulnerable to frauds; then placing sufficient controls in the internal control system will be facilitated. This would also reduce the cost and time involved in forensic assignments, thereby making them more practicable and accessible for any business organization.

Based on this, following research questions arose during the process-
1) Can published financial statements give the indications about the possibility of fraud?
2) Can financial ratios help in identifying the vulnerability of companies to fraud risk which in turn would lead to highlighting the need of forensic accounting?
3) Will the fraud symptoms be different for different sectors?
4) Should the stakeholders and particularly the investors consider fraud risk as a deciding factor before establishing relationship with a company?
5) Can the forensic accounting services be made more affordable to its users?
6) Is the awareness about utility of forensic accounting and related services seen in corporate world in India?

3.6 Objectives of the research

The objectives of this research were as follows-

1) To study the published financial statements of various companies in different sectors in order to identify the need of forensic accounting.
2) To identify sector – specific red flags which would bring forth the fraud symptoms at an early stage.
3) To develop a statistical model to study the need of forensic accounting with the help of ratio analysis.
4) To facilitate investors to understand the elements of accounting they should study more closely before investing in a company or sector.

3.7 Scope of the research

1) There are many frauds which cannot be directly traced to financial statements but affect the organization. They are called as ‘Off the books’ frauds. Frauds like kick-backs from vendors, frauds using information technology, stealing cash before it is recognized by an accounting entry, unrecorded sales etc. fall in this category. The research under consideration was restricted to only those financial frauds which have a reflection in the books of accounts and affect the financial statements adversely. ‘Off the books’ frauds form part of the present study, to the extent they impact the financial statements.

2) The research was restricted to only the financial and quantitative indicators of frauds, which can be shown in terms of numbers. There exist many qualitative indicators of possibilities of frauds such as pressure to meet debt covenants or analysts' earnings expectations, weak internal controls, changes in auditors
over accounting or auditing disagreements, meeting earnings targets consistently over an extended time frame projecting high growth rates etc. However these indicators have not been considered for this study.

3) The researcher has categorically excluded few other sectors like banks and other financial institutions, agriculture etc. from the research. Banks and financial institutions have a different financial and reporting framework and as such comparing them with other non-financial sectors would not have been appropriate. Agricultural sector is relatively unorganized in terms of record keeping, financial reporting etc.

4) This research was restricted to public companies as defined by the companies act only. Private companies and other forms of commercial organizations were out of the orbit of this study.

3.8 Limitations of the study

1) The research was restricted to possibilities of financial statement frauds only. Other types of frauds like cyber-crimes, identity theft, payroll frauds etc. outside the purview of this study.

2) The researcher faced inherent limitation in the study of forensic accounting as no much research exists in this area in Indian context.

3) The researcher has used averaging technique to attain a measure of central tendency for all the ratios of all companies across the time span under consideration. Due to this, the year on year trends of each company could not be observed.

4) The researcher was more interested in observing the behavior of the indicator ratios of an entire sector. As such, year-on-year study of individual companies could not be taken up in this research.

5) The secondary data used for analysis was not available for an entire span of ten years in case of 5-10% companies due to either closure of such companies or limitations of the database. However, to avoid the aberrations in data, all the variables were averaged and then utilized for data analysis to arrive at rational findings.
3.9 Hypotheses
1) Properly identified red flags or fraud indicators based on financial data lead to early diagnosis of fraud symptoms.
2) Properly identified red flags or fraud indicators based on financial data pinpoint the need of forensic accounting as a management control system for selected areas of transactions instead of implementation of full-fledged forensic accounting system.

3.10 Research Design
The researcher has used ratio analysis as a tool to dig out the indicators prompting the need of forensic accounting. The research is quantitative in the sense that it has explained the fraud symptoms in terms of numbers with the help of ratio analysis and the necessary statistical tests. Quantitative research is a methodology followed in case of the research where the aim is to quantify a phenomenon. Quantitative analysis is easy to interpret for the users of the financial statements and as such has great value and utility.

The research design is based on a strong review of literature. The present study indicated sector-wise outliers in financial data, showing thereby the fraud symptoms and the consequent need of forensic accounting.

3.10.1 Time frame for the study
Every research is based on some time frame and the research is bound by the same. This time frame demarcates the boundary of data selection and allows the researcher to focus on a particular time frame for the research.

Researcher, in the present study, chose the time frame of ten years between 2004 and 2013 for the secondary data used in the research under consideration.

Primary data was collected from middle and senior finance executives in order to understand their awareness about the field of forensic accounting, fraud risk mechanism and control etc.
3.11 Data Collection Methods-

Data collected was a mix of primary and secondary data.

3.11.1 Collection of Primary Data

Primary data was important since it meant getting the first hand opinions from industry practitioners about their ways of coping with the fraud risk and about the perceived need of forensic accounting. Researcher was keen on checking if identification of fraud symptoms is useful to managers of corporate world. Researcher was also curious to check if any formal structure of fraud risk assessment exists in organizations.

The primary data collection and analysis was not aimed at proving the hypotheses of the present study. Primary data was essential only to get a feel about the awareness about the field of forensic accounting and views of the finance executives on the same. Thus, the study was restricted to 18 executives only. The said executives from public companies were approached and a semi-structured questionnaire was executed to seek information from them. Same questions were asked to all the respondents. Distribution of questionnaire was done on the basis of suitability, mostly by personal contacts, emails. In some cases, telephonic conversations were used. Questions asked were a mix of open ended and closed ended questions. Semi-structured questionnaire gave enough flexibility to respondents to share their views without being bound by multiple choice answers for all the questions. At the same time, it ensured that the discussions are kept on track.

Fraud is a very sensitive topic and hence, it was absolutely essential that no questions relating to respondent’s employer or his work be asked. Hence, the researcher categorically ensured that all the questions are generic in nature, allowing the respondents to answer without any hesitation or apprehensions about disclosure of sensitive information.

3.11.2 Collection of Secondary Data

a) Professional Database-

1) The Secondary Data required for the present study was collected through a professionally managed database called “Capitaline TP”. Raw data was
downloaded with appropriate query posed to the database in order to extract a wide dataset. The data was wide in terms of time frame and also in terms of number of companies and sectors involved in this study.

2) Private companies were removed from this data since the scope of the research is restricted to published financial statements only, thereby restricting the study to public companies only. A sample of 2595 companies across 17 different sectors was finalized as the input for the present study.

b) Other Sources of Information
Apart from this, secondary data was also collected from many sources like websites; ACFE reports, survey documents published by chartered accountancy firms; reports from regulatory authorities like revenue department, SFIO etc.; books on forensic accounting like forensic accounting for dummies, publication by Bombay chartered accountants association; doctoral theses on related topics etc.

All these sources helped the researcher in understanding the subject more intensely and the way forward was then relatively easier.

3.11.3 Pilot Study
A pilot study was conducted using the financial data of ten companies each from three sectors namely Trading, Shipping and Diversified. Data for ten years from 2004 and 2013 were picked up for the purpose of pilot study. Eighteen ratios were calculated. Then they were averaged and processed further with the help of ‘Descriptive Analysis using Mean and Standard Deviation’ and Correlation analysis.

Descriptive analysis gave broader ratios whereas correlation gave more pointed ratios as red flags. The findings showed that using mean and standard deviation as tools to detect indicators showing fraud possibilities, were very effective. As such, these tools could be used for the entire dataset also.

Results of pilot study showed that when worked with bigger data, findings can be more fruitful. As such, the researcher decided to take 17 major sectors representing a wider section of the Indian economy on the whole.
Also it was felt that more advanced techniques were required to be used in order to
device a statistical model for early detection of red flags. Researcher found out from
the literature review that, ‘Z scores’ are superior in identifying fraud indicators as
compared to using only mean and standard deviation. Thus, final data analysis has
been performed using ‘Z Scores’ and subsequently logistic regression and chi square
test of contingency were performed on the secondary data, as deemed fit.

3.11.4 Reliability and Validity of the questionnaire used for primary data
collection

Reliability of the questionnaire ensures that the questions asked measure exactly those
things that they are supposed to measure. Reliability of the questionnaire was assessed
using test-retest method for those questions which involved ranking of the given
options by the respondents. Remaining questions were either open ended or had very
clear options to select for the respondents. As such, reliability study was not really
essential for those questions.

Test-retest is the most direct method of attesting the reliability of questions. In this
method, the consistency of responses to a scale or instrument is verified by
administering the scale or instrument to the same sample of respondents on two
different occasions. Then, a correlation coefficient is computed on the two sets of
responses. Correlation coefficient value of more than 0.8 is considered to be an
indication of reliable question.

Table 3.1- Reliability checking for questionnaire used in primary data collection

<table>
<thead>
<tr>
<th>Ques. No.</th>
<th>Pair</th>
<th>Correlation Coefficient value</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q.2a</td>
<td>Employee Frauds (Test) ↔ Employee Frauds (Retest)</td>
<td>0.991</td>
<td>Reliability supported</td>
</tr>
<tr>
<td>Q.2b</td>
<td>Cybercrimes (Test) ↔ Cybercrimes (Retest)</td>
<td>0.948</td>
<td>Reliability supported</td>
</tr>
<tr>
<td>Q.2c</td>
<td>Financial statement frauds (Test) ↔ Financial statement frauds (Retest)</td>
<td>0.992</td>
<td>Reliability supported</td>
</tr>
<tr>
<td>Q.2d</td>
<td>Physical Theft (Test) ↔ Physical Theft (Retest)</td>
<td>0.964</td>
<td>Reliability supported</td>
</tr>
<tr>
<td>Q.2e</td>
<td>Misplaced Identity (Test) ↔ Misplaced Identity (Retest)</td>
<td>0.979</td>
<td>Reliability supported</td>
</tr>
<tr>
<td>Ques. No.</td>
<td>Pair</td>
<td>Correlation Coefficient value</td>
<td>Result</td>
</tr>
<tr>
<td>----------</td>
<td>--------------------------------</td>
<td>---------------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>Q.8a</td>
<td>Frauds against organization (Test) ↔ Frauds against organization (Retest)</td>
<td>0.990</td>
<td>Reliability supported</td>
</tr>
<tr>
<td>Q.8b</td>
<td>Frauds for personal gain (Test) ↔ Frauds for personal gain (Retest)</td>
<td>0.765</td>
<td>Reliability marginally missed</td>
</tr>
<tr>
<td>Q.8c</td>
<td>Frauds for organization (Test) ↔ Frauds for organization (Retest)</td>
<td>0.884</td>
<td>Reliability supported</td>
</tr>
<tr>
<td>Q.9a</td>
<td>Tips from employees (Test) ↔ Tips from employees (Retest)</td>
<td>0.926</td>
<td>Reliability supported</td>
</tr>
<tr>
<td>Q.9b</td>
<td>Past financial ratios (Test) ↔ Past financial ratios (Retest)</td>
<td>0.801</td>
<td>Reliability supported</td>
</tr>
<tr>
<td>Q.9c</td>
<td>Important documents (Test) ↔ Important documents (Retest)</td>
<td>0.883</td>
<td>Reliability supported</td>
</tr>
<tr>
<td>Q.10a</td>
<td>Ratio analysis (Test) ↔ Ratio Analysis (Retest)</td>
<td>0.934</td>
<td>Reliability supported</td>
</tr>
<tr>
<td>Q.10b</td>
<td>Whistle blower (Test) ↔ Whistle blower (Retest)</td>
<td>0.810</td>
<td>Reliability supported</td>
</tr>
<tr>
<td>Q.10c</td>
<td>Audits (Test) ↔ Audits (Retest)</td>
<td>0.550</td>
<td>Reliability violated</td>
</tr>
<tr>
<td>Q.12a</td>
<td>Financial statement frauds (Test) ↔ Financial statement frauds (Retest)</td>
<td>0.708</td>
<td>Reliability marginally missed</td>
</tr>
<tr>
<td>Q.12b</td>
<td>Cybercrimes (Test) ↔ Cybercrimes (Retest)</td>
<td>1.000</td>
<td>Reliability supported</td>
</tr>
<tr>
<td>Q.12c</td>
<td>Employee led frauds (Test) ↔ Employee led frauds (Retest)</td>
<td>0.904</td>
<td>Reliability supported</td>
</tr>
<tr>
<td>Q.13a</td>
<td>Internal controls (Test) ↔ Internal controls (Retest)</td>
<td>0.785</td>
<td>Reliability marginally missed</td>
</tr>
<tr>
<td>Q.13b</td>
<td>Statutory audit (Test) ↔ Statutory audit (Retest)</td>
<td>0.979</td>
<td>Reliability supported</td>
</tr>
<tr>
<td>Q.13c</td>
<td>Fraud experts (Test) ↔ Fraud experts (Retest)</td>
<td>0.711</td>
<td>Reliability marginally missed</td>
</tr>
<tr>
<td>Q.13d</td>
<td>Whistle blowing (Test) ↔ Whistle blowing (Retest)</td>
<td>0.737</td>
<td>Reliability marginally missed</td>
</tr>
<tr>
<td>Q.14a</td>
<td>Employee training (Test) ↔ Employee training (Retest)</td>
<td>0.809</td>
<td>Reliability supported</td>
</tr>
<tr>
<td>Q.14b</td>
<td>Fraud experts (Test) ↔ Fraud experts (Retest)</td>
<td>0.953</td>
<td>Reliability supported</td>
</tr>
<tr>
<td>Q.14c</td>
<td>Data analytics (Test) ↔ Data analytics (Retest)</td>
<td>0.821</td>
<td>Reliability supported</td>
</tr>
</tbody>
</table>

From the above table, it was seen that reliability was supported for all variables except one.

**Face Validity**

The questionnaire was designed after a detailed discussion with the academicians, industry practitioners, chartered accountants and company secretaries. The researcher
considered all the valuable inputs given by these experts and included them in the questionnaire.

3.12 Sampling Method
Sample selection is a crucial process and determines the accuracy of results and strength of the research in establishing facts, to a great extent. Researcher in the present study, has selected the sample for primary and secondary data in the following manner:

3.12.1 For Secondary Data
Researcher has used secondary data with a wide time frame due to following factors-
1) Data collected from financial statements is most reliable, since it is a formal communication from the company and is authentic to the extent of statutory audit procedures.
2) Financial statements are relied upon by stakeholders for any decision making.
3) Financial statements are the most common source of information with ease of access to stakeholders and as such easy to use for identification of red flags.
4) It was found from the review of literature that using financial statements as an input for fraud detection techniques proves to be very effective. Thus, the researcher used it for identifying indicators of vulnerability of companies to fraud risk.

3.12.1.1 Defining the universe
The entire Indian economy was the universe of this study. Since the study was about various sectors of the economy, it was not bound by geographical limitations but covered the wider spectrum of the country as a whole.

3.12.1.2 Sectors selected for the study
A total of 17 sectors had been chosen for the study. These sectors covered the majority of companies of the economy and the data between the ten years’ span ranging from 2004 to 2013 was collected.
Following is the list of sectors chosen-
Table 3.2 – List of Sectors and number of companies considered for the present study

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Name of the sector</th>
<th>Number of Companies</th>
<th>Industries Included</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Automobiles</td>
<td>194</td>
<td>Auto Ancillaries, Trucks, Cars, Two wheelers etc.</td>
</tr>
<tr>
<td>2</td>
<td>Chemicals</td>
<td>115</td>
<td>Organic and inorganic, alcohol and specialty – Small, medium and large companies</td>
</tr>
<tr>
<td>3</td>
<td>Coal and Mining</td>
<td>116</td>
<td>Coke, Copper, Copper alloys, minerals, zinc, Mining and Minerals, Refineries, Refractories, Oil Drilling and Crude Oil</td>
</tr>
<tr>
<td>4</td>
<td>Construction Materials</td>
<td>102</td>
<td>Cement, ceramics, cement products, laminates, paints etc.</td>
</tr>
<tr>
<td>5</td>
<td>Construction</td>
<td>334</td>
<td>Housing and commercial – small, medium and large companies</td>
</tr>
<tr>
<td>6</td>
<td>Consumer Goods</td>
<td>119</td>
<td>Personal care, domestic appliances, decoratives etc.</td>
</tr>
<tr>
<td>7</td>
<td>Diversified</td>
<td>18</td>
<td>All the sundry works not clubbed in any other sector</td>
</tr>
<tr>
<td>8</td>
<td>Electricals and Electronics</td>
<td>144</td>
<td>Electric equipment like gensets, picture tubes and others, electrodes, electronic items etc.</td>
</tr>
<tr>
<td>9</td>
<td>Food</td>
<td>226</td>
<td>Dairy products, food processing – small, medium and large companies</td>
</tr>
<tr>
<td>10</td>
<td>Hotels and Resorts</td>
<td>96</td>
<td>Small, medium and large hotels, recreation and amusement parks, resorts etc.</td>
</tr>
<tr>
<td>11</td>
<td>Information Technology</td>
<td>318</td>
<td>Info-tech / database, ITES, Software companies – small, medium and large, hardware – small, medium and large etc.</td>
</tr>
<tr>
<td>12</td>
<td>Packaging</td>
<td>82</td>
<td>Plastic containers, metallic, BOPP, lamination/processors etc.</td>
</tr>
<tr>
<td>13</td>
<td>Pharmaceuticals</td>
<td>168</td>
<td>Bulk drugs and formulations</td>
</tr>
<tr>
<td>14</td>
<td>Power Generation and Supply</td>
<td>105</td>
<td>Power generation and supply</td>
</tr>
<tr>
<td>15</td>
<td>Shipping</td>
<td>38</td>
<td>Small, Medium and Large companies in shipping as also Repairs and breaking and Ship Building</td>
</tr>
<tr>
<td>16</td>
<td>Steel</td>
<td>168</td>
<td>Small, Medium and Large companies in steel as also companies working in areas like rolling, Sponge Iron, tubes/pipes, wires etc.</td>
</tr>
<tr>
<td>17</td>
<td>Trading</td>
<td>252</td>
<td>Small, Medium and Large companies</td>
</tr>
<tr>
<td></td>
<td><strong>Total</strong></td>
<td><strong>2595</strong></td>
<td></td>
</tr>
</tbody>
</table>
3.12.1.3 Sample Size
A total of 17 sectors from Indian economy were picked up for study. All the companies falling between financial year 2004 and 2013 were picked up as sample. This made the total sample size to be 2595 companies. Private companies have been excluded from the study.

3.12.2 For Primary Data
The semi-structured questionnaire was prepared and circulated to 40 middle and senior level finance executives and 18 respondents could actually be contacted and information was gathered from them. The rate of questionnaires received back was 45% of the total circulation. The collection of the primary data was done through interviews as also getting the information with the help of the semi-structured questionnaire mentioned above. Since the subject of fraud is quite sensitive, the respondents answered the questions only on the condition that their names and other details would remain undisclosed and the data would be used for academic reasons only.

Due to the limitations on time, geographical distance and availability of the respondents for the survey, the interviews were conducted face to face as also telephonically in some cases.

The sampling method used was ‘convenient sampling’ and was ‘random’ in nature. The main objective of these questionnaires-cum-interviews was to get a feel of the level of awareness about fraud risk mitigation and use of forensic accounting techniques in corporate world in India.

3.13 Variables used in Data Analysis in the present study
3.13.1 Variables of the study used for Secondary Data analysis
Secondary data was analyzed using following dependent and independent variables.

3.13.1.1 Dependent Variables
The dependent variable used in the study was the ‘vulnerability to fraud risk’.
3.13.1.2 Independent Variables
Researcher focused on 18 distinct financial ratios. All those ratios have been taken to be independent variables. These ratios were calculated and further processed with the help of the statistical tests used in the research. Data was collected from the 18 financial ratios have been listed below-

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Ratio</th>
<th>Abbreviation used in data analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Current Ratio</td>
<td>Current Ratio</td>
</tr>
<tr>
<td>2</td>
<td>Current Assets / Total Assets</td>
<td>CA / TA</td>
</tr>
<tr>
<td>3</td>
<td>Working Capital / Sales</td>
<td>WC / Sales</td>
</tr>
<tr>
<td>4</td>
<td>Inventory / Total Current Assets</td>
<td>Inventory / TCA</td>
</tr>
<tr>
<td>5</td>
<td>Debtors / Total Current Assets</td>
<td>Drs / TCA</td>
</tr>
<tr>
<td>6</td>
<td>Cash / Sales</td>
<td>Cash / Sales</td>
</tr>
<tr>
<td>7</td>
<td>Sales / Debtors</td>
<td>Sales / Drs</td>
</tr>
<tr>
<td>8</td>
<td>Sales / Inventory</td>
<td>Sales / Inv</td>
</tr>
<tr>
<td>9</td>
<td>Sales / Working Capital</td>
<td>Sales / WC</td>
</tr>
<tr>
<td>10</td>
<td>Sales / Total Assets</td>
<td>Sales / TA</td>
</tr>
<tr>
<td>11</td>
<td>Debt / Equity Ratio</td>
<td>D /E</td>
</tr>
<tr>
<td>12</td>
<td>Total Debt / Total Assets</td>
<td>TD / TA</td>
</tr>
<tr>
<td>13</td>
<td>Working Capital / Total Assets</td>
<td>WC / TA</td>
</tr>
<tr>
<td>14</td>
<td>Gross Profit / Total Assets</td>
<td>GP / TA</td>
</tr>
<tr>
<td>15</td>
<td>Net Profit / Total Assets</td>
<td>NP / TA</td>
</tr>
<tr>
<td>16</td>
<td>Net Profit Ratio</td>
<td>NP Ratio</td>
</tr>
<tr>
<td>17</td>
<td>Inventory / Sales</td>
<td>Inv / Sales</td>
</tr>
<tr>
<td>18</td>
<td>Gross Profit Ratio</td>
<td>GP Ratio</td>
</tr>
</tbody>
</table>

3.13.2 Variables of the study used for Primary Data Analysis
Variables for primary data involved various key concepts like fraud mitigation, detection techniques, awareness about forensic accounting etc.
3.14 Measures of financial performance: Ratios and Explanation about each of them

3.14.1 Meaning of Financial Ratios and Ratio Analysis

(A) Financial Ratios

A ratio is a relationship between two variables expressed in terms of one variable divided by the other. Financial ratios are mathematical comparisons of financial statement accounts or groups. Financial management incorporates the study of the relationships between various financial variables in such a manner that it helps investors, creditors, and all the stakeholders in understanding the financial health of a business concern.

Financial ratios are the analytical tools which are most commonly used for the analysis of the financial condition of any business.

Ratios are easy to comprehend and simple in calculation. They help to analyze the financial data of companies in many ways:
- Inter firm comparison for a financial year
- Inter firm comparison between two or more financial years
- Intra firm comparison for a financial year
- Intra firm comparison between two or more financial years
- Establishing relationship between several variables in one or more financial years
- Inter firm comparison of different companies in different industries

(B) Financial Ratio Analysis

(ACFE, 1999) in its literature on frauds stated that ratio analysis is a technique of determining the relationship between two different financial variables. (Leeds, 2013) in his book stated that financial ratios enable the analyst to use and compare companies of varied sizes to each other directly. Thus, ratios remove the issue of company size as an aspect of consideration.

This system of financial statement analysis can be very useful in detecting red flags for a fraud examination. Any significant change in the values of ratios from one year
to the other, or over a period of many years might mean a financial difficulty or misrepresentation of financial data or fraud or a genuine error. An aberration in the trend of a particular ratio studied over years should pop up a red flag for financial data users.

(ASOSAI) in its notes on auditing mentioned that red flag would indicate a need to dig deep into a financial transaction as also into its record keeping to know if the aberration is incidental or intentional. For example, increase in sales not supported by corresponding increase in sales commission increasing trend of debtors, increase in inventory not in line with corresponding increase in sales etc. are a few examples of a red flag, indicating an anomaly needing special attention by accountants as also auditors. Ratios help in such anomaly detection to a large extent.

3.14.2 Rationale behind choosing the ratios used for the present study
Many researchers like (Persons, 1995), (Spathis, 2002) have studied the utility of various ratios in predicting future behaviour of financial data of a company in the literature available till date. Ratio analysis is widely used and most common form of analysis due to its benefits like ease of calculation, simple to understand etc.

However, researchers like (Kaminski, 2004) have mentioned that these ratios have limited ability in detecting frauds. Researchers like (Green & Choi, 1997), (Panigrahi, 2006) on the other hand, mentioned that data mining tools when used along with ratios, do enhance the ability of ratios to predict and detect fraud possibilities. Researcher in the present study has thus used various financial ratios and have coupled them up with statistical tools namely outlier analysis also called as anomaly detection, standard binary logistic regression and chi square test of contingency as and when needed (Sharma & Panigrahi, 2012). Most of the research conducted by using financial ratios till date has been in the field of fraud detection and not much in prevention area.

The researcher thus, used these ratios for fraud prevention and early detection of frauds, using financial statements. A total of 18 financial ratios which were considered useful, based on the previous literature and knowledge of the researcher in
the area of auditing and financial management. A detailed explanation about all these ratios is given in the next section of this chapter.

3.14.3 Meaning and formula used for each ratio considered for the present study
Twenty seven variables like share capital, secured and unsecured loans, net sales, interest, gross profit, reported net profit etc. were extracted from the Capitaline database and 18 financial ratios were calculated thereafter, with the help of the extracted variables.

The financial ratios used for the “Data Processing” have been explained hereunder-

(1) **Current Ratio**-
Current ratio is generally used in analysis, in order to evaluate the short term liquidity and solvency position of a firm. It explains the availability of cash and cash equivalent in a firm as against the liabilities it holds. It is calculated as “Current Assets / Current Liabilities”.

This ratio is one of the key ratios which can indicate manipulations in financial statements, if any. Drastic change in current ratio would mean the need to go into detailed analysis of financial statements and investigate the causes of the manipulations. Concealment of liability will show a more favourable ratio, while showing false debtors or inventory will increase the ratio artificially. Cheque tampering resulting in decrease in cash, thereby showing a decline in the current ratio or significant increase or decrease in current ratio can be a few red flags for this ratio.

(2) **Current Assets to Total Assets**-
This ratio indicates the proportion of current assets in relation to total assets of a firm. (Persons, 1995) has found out that the proportion of current assets in relation to total assets is found to be relatively higher in case of fraud firms when compared with non-fraud firms. Thus, studying this ratio was strategically relevant in the present study.
(3) **Working Capital to Sales** –
This is a liquidity ratio. It indicates the efficiency with which a company has used its working capital to generate sales. More the sales, more will be the profit. And more profit means the company is efficient in using the same working capital over and over again. Companies committing financial statement frauds will usually have earnings growth not correspondingly supported by a corresponding growth in cash flows. This means their current assets other than cash will inflate or current liabilities will decrease in order to try to match up the figures. Thus this ratio becomes extremely crucial to dig out for a red flag. (ACFE Conference, 2012).

(4) **Inventory / Total Current Assets**-
The ratio indicates the money invested in inventory as a proportion of total current assets. It indicates the liquidity position of a company also. Higher ratio is indication of inefficient inventory management or large obsolete inventory in total inventory of company (Persons, 1995).

Certain items of financial statements involve an element of subjectivity due to the methods by which they are measured in monetary terms. Debtors and Inventory are primarily the items coming under this category. Various methods of Inventory valuation, subjectivity in deciding reserve for bad debts etc. leads to opportunity of misstatement in the components of current assets. A study by (Vanasco, 1998) stated that auditors treat inventory, cash defalcation and conflict of interest as the items with very high possibilities of fraud. Kaminski et al. (2004) had chosen this ratio of inventory to current assets as one of the possible predictors of fraud.

(5) **Debtors / Total Current Assets** -
Many researchers have proposed that items in current assets such as account receivables and inventories are usually considered to be more prone to manipulation. Fraudulent companies have been found to be manipulating these items more often than others (Spathis, 2002). Thus, this ratio was picked up for the present study.
(6) **Cash / Sales** –

This ratio depicts the effectiveness of credit policies of a company. It states the amount of cash needed as buffer at any point of time. It is said to be a good measure of asset allocation and liquidity. Yap et al. (2010) used this ratio to predict failure of a company.

(7) **Sales / Debtors** –

(Persons, 1995), (Schilit, 1993) and (Green & Choi, 1991) suggest that management may manipulate receivables. Showing additional receivables would increase the current assets and current ratio will improve as an effect. Fake credit sales recorded by a firm in order to show more revenue would also result in increasing the amount of receivables. Apart from this; aging schedule of debtors, timing of recording a transaction, provision for bad debts, write off of bad debts are all likely to be manipulated. These relations thus become essential in this research.

(8) **Sales / Inventory** –

This ratio depicts how quickly a firm has used its inventory to generate the revenue. Purpose of this ratio is to measure the liquidity of a company. A high ‘sales to inventory’ ratio may mean either strong sales or inefficient buying or inadequate inventory. If the ratio is too low, this may mean either poor sales or excess inventory not commensurate with the rate of sales.

(Persons, 1995) in his study found that one of the most obvious factors associated with fraudulent financial statements is ‘Inventories’. Some companies in USA have been found reporting its inventory not at its actual value and recording obsolete inventory (Lieu and Yang, 2008). (Loebbecke et al., 1989) found in their study that inventory and account receivables were involved in 12-14 percent of falsified financial reporting respectively. Thus, this ratio was very crucial and was considered by the researcher.

(9) **Sales / Working Capital** -

This is an efficiency ratio. This ratio measures asset utilization of a firm. It helps the stakeholders to understand how much cash is needed to generate a certain level of sales. It can throw light on company's need to raise additional funds in order to
increase the sales. Ratios that are lower than the industry average are generally suspicious and open for investigation.

(10) Sales / Total Assets-
(Albrecht, Albrecht, Albrecht, & Zimbelman, 2011) stated that assets turnover and working capital turnover ratio can help in detecting frauds related to any of the components of the current assets such as inventory, debtors, cash etc.

This ratio was also found to be relevant in prior research by (Persons, 1995) and (Loebbecke et al., 1989). (Altman, 1968) used this ratio for prediction of financial difficulties and thus, considered for present study.

(11) Debt to Equity –
It indicates the amount of debt borrowed by a firm when compared to its equity. Higher the ratio, more is the burden on the owners on account of repayment of debt. It also means that creditors have shouldered more responsibility of funding the business. (Spathis, 2002) stated that higher leverage is usually associated with higher potential for violations for loan agreements and lesser ability to obtain additional capital through borrowing. Pressure from shareholders to maintain delicate balance between equity and debt may tempt the management to misstate this ratio and as such considering it for this research was felt necessary by the researcher.

(12) Total Debt to Total Assets-
(Persons, 1995) and (Spathis, 2002) have stated that higher level of debt may lead to higher possibility of fraud.

(13) Working Capital to Total Assets-
(Spathis, 2002) and Loebbecke et al. (1989) had used this ratio. Lower the ratio higher is the likelihood of fraud and misstatement.

(14) & (15) Gross Profit to Total Assets and Net Profit to Total Assets-
(Spathis, 2002) brought up the point as to whether higher or lower margins can give indications about fraud possibility and included ratio of Gross Profit to Total Assets in his study. (Loebbecke et al., 1989) also used these ratios in his study. (Mehta Ujal et
al., 2012) confirmed that net profit to total assets have an increased possibility of predicting fraudulent firm and as such these ratios were included in the present study.

(16) **Net Profit Ratio** –
Net Profit ratio is calculated as \[(\text{Net Profit} / \text{Sales}) \times 100\]. This ratio helps to understand the profitability of a business for a given financial year. Companies with low profitability tend to manipulate figures more than healthy companies.

(17) **Inventory / Sales** –
This ratio explains the variations in the level of investment. It helps to predict cash flow problems related to your business's inventory at an early stage. An increase in this ratio means that either the sales are dropping or the investment in inventory is growing more speedily than the growth in sales. Aspects like goods on sale or return basis, cash or accrual method of recording sales, timing of recording sales etc. are very crucial and sensitive areas as far as misstatement is concerned.

Mehta Ujal et al. (2012) proved that firms with higher inventories to sales have a higher risk of being categorized as a firm with fraudulent financial statements.

(18) **Gross Profit Ratio**
Gross Profit ratio is calculated as \[(\text{Gross Profit} / \text{Sales}) \times 100\]. Researchers like (Salama, 2014), Spathis (2002) etc. have used this ratio as a possible indicator for fraud detection and even company failure. (Fanning and Cogger, 1998) mentioned that discrepancies can be found in gross profit. Sales for a year may not with the corresponding cost of goods sold, thereby increasing gross margin, net profit and making the balance sheet look healthier. Thus, the researcher considered this ratio for the research.

3.14.4 Assumptions in calculation of the ratios
(1) Sales have been taken as net of excise. Excise duty is to be paid on the value of goods produced and as such will not affect the ratios in any manner. Sales figure after deducting excise will reflect the true sale price and as such would be better at indicating red flag. Prof. L.C. Gupta in his research on industrial sickness conducted in 1983, mentions four important profitability ratios as
indicators showing industrial sickness. They are as follows- EBIT / Sales net of excise, DCF / Sales net of excise, EBIT / Total Assets plus accumulated depreciation and DCF / Total Assets plus accumulated depreciation. He took ‘Sales net of excise’ for all calculations.

(Bodhanwala, 2009) stated in this regard that sales turnover always includes a component of excise duty and its possible that this ratio may be largely affected due to any change in excise duty. In case of companies falling in the regime of higher excise duty, gross sales amount may show an increase in sales over previous year, whereas the fact may be that sales have gone up as a result of excise duty. As such taking “Sales net of excise duty” is always advisable.

(2) Ratios namely “Sales to Inventory” “Inventory to Sales” and “Sales to Working Capital” and ‘Working Capital to Sales” have been included purposefully in order to get a different outlook of the same two variables. Since inventory and working capital are the most sensitive areas in relation to misstatements, researcher found it essential to include both the relationship of same two variables.

(3) “Total Assets” are defined as total of current assets and fixed assets of a business for a given financial year. The database used by the researcher for extraction of financial data gave the figure of total assets net of current liabilities and provisions. The researcher has however taken the total assets as per the definition of the term. Thus “Current liabilities and Provisions” were added back to the given figure of “Total assets” in order to arrive at the logical figure.

3.15 Data Processing

3.15.1 Processing of Secondary Data

The secondary data gathered by the researcher was processed in five steps-

i) The secondary data collected and duly sorted, was further processed by sorting company wise and then chronologically. The data for the time period between 2004 and 2013 was collected. It has been observed that taking data for a
longer duration gives better results. This data was now ready for being used for calculation of financial ratios.

ii) Microsoft Excel was used to calculate all 18 financial ratios across 17 sectors for all the years between financial years 2004-2013. Non divisible or redundant figures were eliminated at this stage in order to ensure accuracy and precision in calculation of averages, which was the next step in the data processing. Elimination of redundant figures removed the variations in data and helped in getting more informative and relevant figures.

The “yearly” financial ratios so calculated, were further averaged to get one figure of each of the 18 ratios for every company. Average is said to be the measure of central tendency and as such is effective when inter-firm comparison within a sector is performed.

iii) Statistical Package for the Social Sciences (IBM-SPSS) 21.0 was used for advanced data analysis. The averaged ratios as arrived at in the step (i) above, were used to generate Z scores for each sector. Z scores enabled the comparison between various firms by creating a common base for the same.

All the ratios beyond a threshold of ‘Plus or Minus 2’, written in statistical terms as ‘± 2’ were treated as outliers. In financial statements, an outlier arouses suspicion about possibilities of some irregularities or anomalies in accounting data leading to risks of frauds or errors.

A company has been classified as a “Red Flag” company if it has one or more outliers in the set of 18 ratios calculated. Red Flag Company has been denoted by ‘1’. A company which does not have even one outlier ratio has been termed as a “Green Flag” company and has been denoted by ‘0’.

iv) These Z Scores and the classification as red and green flag companies was used for the advanced data processing performed by the researcher with the help of statistical technique named as “Logistic Regression”. A regression equation has been formulated for 12 sectors.

v) Logistic Regression could not be used for the remaining 5 sectors out of 17 sectors, due to redundancies. Researcher performed ‘Chi Square Test of Contingency’ for those five sectors instead of ‘Logistic Regression’.
Data of five sectors namely Automobiles, Construction Materials, Diversified, Hotels and Resorts and Shipping was processed with the help chi square test which studied the relationship between an individual financial ratio and vulnerability to fraud risk.

For Chi Square test, an additional working had to be done. All the variables must be translated into nominal scale for the chi square test to function effectively. (Zikmund, 2003) stated that nominal scale is best suitable for those cases where the numbers or letters allotted to objects act as tags for categorizing data.

All the z scores were further converted in the format of ‘0’ and ‘1’ depending upon whether a score indicated outlier or not. Companies falling beyond the ‘± 2’ for a particular ratio were called as outliers for that ratio and others were called as inliers for that ratio. All the outlier ratios of all companies were termed as “1” and all the inliers were called as “0”. These inliers and outliers were studied with red flag and green flag companies identified at the time of anomaly detection, already mentioned in (ii) above.

3.15.2 Processing of Primary Data

The primary data was processed and further analyzed using Friedman’s chi square test and pie diagrams in percentage, depending on the type of questions asked and the format in which the answers were expected. The semi-structured questionnaire could be decomposed as follows-

1) Friedman’s test was used when the questions asked were closed ended and in answers were expected in the form of rating or ranking.

2) In case of other closed ended questions, percentage pie diagram was used.

3) There were three open ended questions in the questionnaire. The data processing in their case was initiated by deciphering the answers. The responses were noted and coded and classified. Further analysis was performed using percentage pie diagram.
3.16 Conclusion

The researcher has used secondary data of financial statements of selected companies in order to conduct the study. The aim of the researcher was to use such data as is freely available to everyone concerned. Easy access to input data would improve the utility and the ease with which, such research can be operationalized by the users of the same. Use of advanced statistical tools allowed the researcher to study various facets of the research problem.

Vulnerability to fraud risk generates the need to employ fraud control mechanism for a company. Preventive measures for fraud can be implemented effectively only when supported by knowledge and expertise of forensic accounting principles and practices. It is believed that early warning signals of fraud show the path towards fraud prevention and also help in fraud detection at an early stage. Report by ACFE in 2014 stated that financial statement frauds like payroll frauds, financial statement fraud, reimbursement of expenses frauds etc. had a median duration of 24 months. This highlights the importance of early detection of frauds as also the need of forensic accounting.

The researcher performed the data analysis for the sectors selected for this study, after having designed a robust research design. Suitable statistical tools have been used to arrive at findings and conclusions have been drawn on that basis.
4.1 Introduction
Research is said to be a movement from known to unknown. The goal of any research is to make an original contribution to the present knowledge base. Data analysis can be explained as the method of processing the raw data with the help of statistical tools, in order to establish meaningful relationship between one or more dependent and independent variables. Data analysis helps researcher in studying the sample selected for the study in the light of logical and statistical tools which give a meaning to the raw data.

(Cooper, Schindler, & Sharma, 2012) defined data analysis as “Editing and reducing accumulated data to a manageable size, developing summaries, looking for patterns and applying statistical techniques”. It is the next logical step after data collection.

Data analysis helps researcher in communicating the findings in a simple yet insightful language. Information can be communicated well and understood easily by its users, when presented using graphs, charts, diagrams, equations etc. Many statistical tools and techniques have been empirically used for data analysis.

Researcher in the present study has used data mining technique for analysis of data. (Cooper et al., 2012) in their book have defined data mining as a process of discerning knowledge from database stored in data marts or data warehouses. Purpose of data mining is to identify valid, unique, suitable and logical patterns in the data. Data warehouse is an electronic repository for the database that organizes large volumes of data into categories to facilitate retrieval, interpretation and sorting by the end users. It is said that data in warehouse was once primary data which was collected and prepared for a specific use. But when such data is mined by researcher, it becomes secondary data. Thus, mining can be referred to as the bridge between primary and secondary data. This analogy fits very well to the financial data used in the present study. The financial statements of companies selected by researcher were once primary when prepared; but are treated as secondary when used by researcher. In this sense, the financial data emanating from financial statements is secondary in nature. (Cooper et al., 2012) explained that data mining is comprised of four steps.
Table 4.1 Steps in Data Mining

<table>
<thead>
<tr>
<th>Stage</th>
<th>Meaning</th>
<th>Relevance for the topic at hand</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample</td>
<td>Extracting data from the data warehouse</td>
<td>2595 companies across 17 sectors were selected for study</td>
</tr>
<tr>
<td>Explore</td>
<td>Ascertaining the relationships between variables</td>
<td>Financial ratios were calculated financial years 2004 and 2013 and then averaged.</td>
</tr>
<tr>
<td>Modify</td>
<td>Transforming data into more meaningful form</td>
<td>Z scores</td>
</tr>
<tr>
<td>Model</td>
<td>Preparation of model to explain relationships and check the predictive ability of the model</td>
<td>Model development using logistic regression for 12 sectors and testing done with the help of test like Hosmer Lemeshow test, omnibus chi square test etc. Study of relationships between dependent and independent variables using Chi square test of contingency for 5 sectors</td>
</tr>
</tbody>
</table>

Researcher in the present study has used logistic regression and chi square test of contingency to analyse the secondary data. Primary data was processed using ‘Friedman’s Chi Square Test’ and other suitable statistical tools.

4.2 Chapter Scheme

The chapter is divided into three parts. Each section depicts the details of the entire data. These sections have been designed in following manner-

**Section I and II - Sector wise Assessment of vulnerability to fraud risk**

The assessment of vulnerability to fraud risk for the sectors under study was done with the help of financial analysis of financial ratios. Logistic regression was performed on 12 sectors (Section I) and 5 sectors were processed using chi square test of contingency (Section II).

**Section III – Analysis of awareness about forensic accounting and financial statement frauds in Indian corporate sector using primary data**

This section includes the primary data analysis based on a survey of 18 middle and senior level finance executives of few companies representing various sectors under study. The aim of the study was to know about the awareness of the forensic accounting amongst the finance executives and to understand their ways of coping with fraud risk in general and about financial statement frauds in particular.
4.3 **Statistical Tests and Tools used in the present study**

4.3.1 **Z Scores**

Z scores are standard scores measured in standard deviation units. With reference to a standard normal curve, z score indicates the number of standard deviations a particular value is above or below mean. Z scores with positive values are above mean and with negative are below mean.

A z-score is a measure of how many standard deviations below or above the population mean a raw score is. A z-score is also known as a **standard score** and it can be placed on a normal distribution curve. Z-scores range from -3 to +3 standard deviations; where -3 would fall to the extreme left of the normal distribution curve, and +3 would fall to the extreme right of the curve.

\[
Z \text{ Score} = \frac{\text{Value} - \text{Mean}}{\text{Standard Deviation}}
\]

Z score normalizes the sampling distribution and allows a meaningful comparison. It is a very effective for large data. Researcher used this tool due to the size of data in the present study.

In this research, z scores have been used to detect outliers in every sector amongst the 18 financial ratios worked out using the financial statement data. These outliers were called as ‘Anomalies in financial data’ in the present study. The objective behind using this statistical tool was to find out early indicators of fraud risk arising in the form of anomalies for each company. These anomalies would give the indications about the fraud symptoms for companies in a sector.

4.3.2 **Standard Binary Logistic Regression – Used for Secondary Data**

**Logistic Regression**

(Hair (Jr.), Black, Babin, & Anderson, 2010) have given a detailed view of various techniques of multivariate analysis in their book. One of them is Logistic regression; which is a special type of regression used to predict and explain a binary categorical variable. The specialty of this regression is that it can encompass nonmetric variables also. Logistic regression can be explained as a study of the relation between a single
non-metric binary dependent variable and a set of metric or non-metric independent variables. An equation of logistic regression:–

\[
\text{Log [odds]} = a + b_1x_1 + b_2x_2 + \ldots + b_nx_n
\]

**Peculiar points about Standard Binary Logistic Regression**-

Logistic regression is its ability to build a logistic model and establishing the relationship between dependent and independent variables. The model gives birth to a logistic regression equation which is the group predicted value.

Each logistic regression equation is expressed in terms of log odds. The odds of an event taking place is defined as the probability that the event occurs divided by the probability that the event does not occur. The log odds are nothing but the natural logarithm of the odds.

In the present study, standard binary logistic regression was performed to study if financial ratios can predict the vulnerability of companies to fraud risk. Findings of binary logistic regression will highlight the need of forensic accounting. Binary Logistic regression involved a categorical dependent variable namely ‘vulnerability to fraud risk’ and a set of predictors namely ‘financial ratios’.

**Model Development**-

One of the objectives of the present study was to develop a statistical model in order to study the need of forensic accounting, with the help of ratio analysis. 18 financial ratios over 17 sectors were analyzed.

The model development was facilitated by the logistic regression and models have been developed for 12 out of 17 sectors selected by researcher for the study. Findings from statistical and financial analysis were brought together and efforts were diverted to present an innovative and unique model to identify the indicators showing the need of forensic accounting.

The model so developed has been presented in the form of a logistic regression equation. Logistic regression distinguished between significant predictors and
insignificant predictors in the context of each of the 12 sectors. Significant predictors in the present study mean those financial ratios which have the capability to predict the vulnerability to fraud risk for companies in a sector. Insignificant predictors were those financial ratios which were relatively less noteworthy and can be ignored for identifying the vulnerability to fraud risk.

An independent variable with a regression coefficient not significantly different from 0 (P>0.10) can be excluded from the logistic regression model. Thus, if probability value (P)<0.10, then the variable is said to have contributed significantly to the prediction of the dependent variable and as such it should be retained in the logistic regression equation.

**Validation and Assessment of the LR model**

(Hair et al., 2010) have further stated the way in which a Logistic regression model is measured for checking the model estimation fit with the value -2 times the log of the likelihood value. The likelihood value is referred to as -2LL or -2 log likelihood. Minimum value for -2LL is zero which means a perfect fit. Thus lower the -2LL value, better is the fit of the model. For this, first a null model is established and then a proposed model on that basis is generated. Usually model fit will improve from null to proposed model and will lead to reduction in the -2LL value.

Assessment of the statistical significance of the -2LL value between 2 models is the final step. If the statistical tests support the significant difference, then we can say that the set of independent variables in the proposed model is significant in improving model estimation fit.

**Parameters in Standard Binary Logistic Regression**

When a logistic regression model is fitted to the given data to see how well it fits in; five important parameters are looked into-
Table 4.2 Parameters in Standard Binary Logistic Regression to study the model fit

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Parameter</th>
<th>Method used to test the parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Strength of the model</td>
<td>Nagelkerke test</td>
</tr>
</tbody>
</table>
| 2       | Significance of the model                     | 1) Omnibus test of model coefficients – to check the Chi square significance  
                                                 | 2) Hosmer and Lemeshow test – To test the goodness of fit                                           |
| 3       | Predictive accuracy of the model               | Both the classification tables – For comparison of the two                                          |
| 4       | Significance of the individual predictors in the model leading to regression equation | Wald test – To study and identify predictors which will find a place in the regression equation. In case of the present study, financial ratios would qualify as the possible significant predictors. The table of variables in the equation has depicted those financial ratios out of 18, which qualify as significant predictors for each sector. |

Each term stated in the above table has been explained hereunder-

Key Terms in Logistic Regression-

1) **Likelihood value**- measure used in LR to represent the lack of predictive fit.

2) **Odds**- The ratio of the probability of an event occurring to the probability of the event not occurring is called as odds ratio. It is used as a measure of the dependent variable in logistic regression.

3) **Wald statistic**- It is a test used in logistic regression for checking the significance of the logistic coefficient.

4) **Hosmer and Lemeshow (HL) test**- It is a classification test where the cases are first divided in ten equal classes. Then the number of actual and predicted events is compared in each class with the chi square statistic. Hosmer and Lemeshow test provides a full proof measure of predictive accuracy that is based on the actual prediction of the dependent variable.

In the present study, logistic regression was performed to confirm whether the 18 financial ratios selected by researcher have an ability to predict the group membership; that is to indicate whether a company is a red flag or a green flag company.
4.3.3 Chi Square Test of Contingency – Used for Secondary data

Chi square test is a statistical technique used to test significance in the analysis of frequency distribution. It studies the relationship of a dependent and an independent variable to each other.

Chi square test was performed to study whether there is a relationship between the financial ratios and the vulnerability of companies to fraud risk. The mean of a chi square distribution is its degrees of freedom. As the degree of freedom increases, the degree of skew decreases.

(Zikmund, 2003) stated that the chi-square contingency test is based on the differences between the observed values and the values that would be expected if the variables were independent. Usually, a cross tabulation of the frequencies is mapped to study and present the two categorical variables.

In the research under consideration, chi square test of contingency has been used to indicate whether a particular financial ratio is a significant indicator of vulnerability to fraud risk or not. This test was used for 5 out of 17 sectors under study.

4.3.4 Friedman’s Chi Square Test– Used for Primary Data

Primary data was collected using a semi-structured questionnaire and interviews were conducted. Researcher performed Friedman’s chi square test on all those questions in the questionnaire, where the answers were to be ranked by respondents. All other questions were studied using percentage pie diagram.

Friedman’s chi square test is a non-parametric test used to test for differences between various groups when the dependent variables are ordinal in nature.