# CHAPTER ONE

## 1.0 Historical background of sport

Sport is an activity involving physical exertion and skill in which an individual or team competes against another or others for entertainment (*oxford dictionary*).The history of sports probably extends as far back as the existence of people as purposive sportive and active beings. Sport has been a useful way for people to increase their mastery of nature and the environment. The history of sport can teach us a great deal about social changes and about the nature of sport itself. Sport seems to involve basic human skills being developed and exercised for their own sake, in parallel with being exercised for their usefulness. It also shows how society has changed its beliefs and therefore there are changes in the rules. Of course, as we go further back in history the dwindling evidence makes the theories of the origins and purposes of sport difficult to support. Nonetheless, its importance in human history is undeniable.

Sports that are at least two and a half thousand years old include hurling (similar to field hockey) in Ireland, harpastum (similar to rugby) in Rome, cuju (similar to association football) in China, and polo in Persia. The Mesoamerican ballgame originated over three thousand years ago.

There are artefacts and structures that suggest that the Chinese engaged in sporting activities as early as 2000 BC. Gymnastics appears to have been a popular sport in China's ancient past. Monuments to the Pharaohs indicate that a number of sports, including wrestling, weightlifting, long jump, swimming, rowing, shooting, fishing and athletics, as well as various kinds of ball games, were well-developed and regulated several thousands of years ago in ancient Egypt. Other Egyptian sports included javelin throwing, high jump, and wrestling. Ancient Persian sports such as the traditional Iranian martial art of Zourkhaneh. Among other sports that originated in Persia are polo and jousting. (wikkipedia.org)

There are many sports whose modern rules were formulated during the mid- or late-19th century. World-wide, this certainly includes many different football codes, lawn bowls, lawn tennis and others. The major impetus for this was the patenting of the world's first lawnmower in 1830. This allowed for the preparation of modern ovals, playing fields, pitches, grass courts, etc.

Perhaps in a reaction to the demands of contemporary life, there have been developments in sport that are best described as post-modern, extreme ironing being a notable example. There is also a move towards adventure sports as a form of escapism,

transcending the routines of life, examples being white water rafting, paragliding, canyoning, base jumping, parkour/free running and more genteelly, orienteering.

## 1.1 Origin of Football

The Origins of football can be found in every corner of the Globe. Civilizations throughout history all played ball games and many of these can be considered forerunners to the modern game, which was established in England in the late nineteenth century. Ball games were first played in Egypt as early as 1800 BCE. However there are claims that suggest ball games were played even earlier in Ancient China, maybe as early as 2500 BCE. There is evidence to show that ball games in Egypt were linked to fertility rites and religious ceremonies. These ball games involved large numbers of people and records indicate that these games were sometimes used as an exercise to till the soil. In addition to the Egyptians; the Greeks, Romans and Chinese (as well as others) all played football games, which are considered to be the forerunners of modern football.

### 1.2.1 INTRODUCTION TO SPAIN LEAGUE

The Primera Division (First Division) of the Liga Nacional de Fútbol Profesional (LFP), commonly known in the English-speaking world as La Liga (pronounced: [la

ˈliɣa], The League), and officially named for sponsorship reasons Liga BBVA (BBVA League) is the top professional association football division of the Spanish football league system. It is contested by 20 teams, with the three lowest placed teams relegated to the Segunda División and replaced by the top three teams in that division. A total of 59 teams have competed in La Liga, nine of which have been crowned champions. Since the 1950s, Real Madrid and Barcelona have dominated the championship. Real Madrid have won the title a record 32 times and Barcelona 21 times. During the 1930s and 1940s and in the last two decades, however, La Liga has seen other champions, including, Atlético Madrid, Athletic Bilbao, Valencia, Real Sociedad, Deportivo, Real Betis, and Sevilla. Real Madrid is the current defending champion.

La Liga is the strongest league in Europe over the past five years, according to UEFA's league coefficient. La Liga is one of the most popular professional sports leagues in the world, with an average attendance of 28,286 for league matches in the 2009–10 seasons. This is the sixth-highest of any domestic professional sports league in the world and the third-highest of any professional association football league in Europe, behind the German Bundesliga and English Premier League. (*www.wikkipedia.org*)

### 1.2.2 INTRODUCTION TO BARCLAY'S PREMIER LEAGUE

The Premier League is an English professional league for association football clubs. At the top of the English football league system, it is the country's primary football competition. Contested by 20 clubs, it operates on a system of promotion and relegation with The Football League. The Premier League is a corporation in which the 20 member clubs act as shareholders. Seasons run from August to May, with teams playing 38 matches each, totaling 380 matches in the season. Most games are played on Saturdays and Sundays, with a few games played during weekday evenings. It was known as the Premiership from 1993 to 2007. It is currently sponsored by Barclays Bank and therefore officially known as the Barclays Premier League.

The competition formed as the FA Premier League on 20 February 1992 following the decision of clubs in the Football League First Division to break away from The Football League, which was originally founded in 1888, and take advantage of a lucrative television rights deal. The Premier League has since become the world's most watched association football league. It is the world's most lucrative football league in terms of revenue, with combined club revenues of over £2 billion in 2008–09. It is ranked first in the Union of European Football Associations (UEFA) coefficients of leagues based on performances in European competitions over the last five years, ahead

of the Spanish La Liga and German Bundesliga. It is the number two football league in terms of average attendance behind the German Bundesliga.

Since 1888, a total of 23 clubs have been crowned champions of the English football system. Of the 45 clubs to have competed since the inception of the Premier League in 1992, five have won the title: Manchester United, Arsenal, Chelsea, Blackburn Rovers and Manchester City. The current champions are Manchester United, who won the title in the 2012–13 seasons. (wikkipedia.org)

### 1.3.1 HISTORICAL BACKGROUND OF SPAIN

The history of Spain is the story of the many different peoples who have lived in the Iberian Peninsula, especially in what today is Spain. It spans from prehistoric Iberia, the middle Ages, through the rise and decline of a global empire, to the recent history of Spain as a member of the European Union. Modern humans entered the Iberian Peninsula about 32,000 years ago. Different populations and cultures followed over the millennia, including the Iberians, the Tartessians, Celts and Celtiberians, Phoenicians, Greeks, Carthaginians, Romans, Suebi and Visigoths.

In 711, the Moors, a Berber and Arab army, invaded and conquered nearly the entire peninsula. During the next 750 years, independent Muslim states were established, and the entire area of Muslim control became known as Al-Andalus. Meanwhile the

Christian kingdoms in the north began the long and slow recovery of the peninsula, a process called the *Reconquista*, which was concluded in 1492 with the fall of Granada. Over time, various petty kingdoms and other states on the peninsula began to coalesce into larger states. Three kingdoms came to dominate Iberia by the 15th century, those being the Kingdom of Portugal, Kingdom of Castile, and Kingdom of Aragon. The dynastic union of the crowns of the Kingdom of Castile and the Kingdom of Aragon in 1492 is seen frequently as the starting point of the unification of the modern kingdom of Spain. Although colloquially and literarily the expression "King of Spain" or "King of the Spain's" was already widespread, and the two crowns were ruled by the same monarch, they retained their individual institutions and identity until 1812, with the enactment of the Spanish Constitution of 1812. Portugal was briefly also ruled by the House of Habsburg with Castile and Aragon in the Iberian Union, but this short-lived arrangement was never popular in Portugal; Portugal after 60 years refused to be ruled by the Habsburg king.

The first voyage of Christopher Columbus to the New World took place in 1492 as well, beginning the development of the Spanish Empire. The Inquisition was established and Jews and Muslims who refused to convert were expelled from the country. For the next century and a half, Spain was the most powerful state in Europe

and the largest overseas empire in the world for three centuries. Spanish literature and fine arts, scholarship and philosophy flourished during the 16th and 17th centuries. Spain established a vast empire in the Americas, stretching from California to Patagonia, and colonies in the western Pacific. Financed in part by the riches pouring in from its colonies, Spain became embroiled in the religiously charged wars and intrigues of Europe, including the Dutch Revolt, the French Wars of Religion, and the Thirty Years War.

Spain's European wars led to economic damage, and the latter part of the 17th century saw a gradual decline of power under an increasingly neglectful and inept Habsburg regime. The decline culminated in the War of Spanish Succession, which ended with the relegation of Spain to the status of a second rate western power, although it remained, with Russia, the leading colonial power of the 18th century. The 18th century saw a new dynasty, the Bourbons, which directed considerable efforts towards the renewal of state institutions, with some success, finishing in a successful involvement in the American War of Independence. However, as the century ended, a reaction set in with the accession of a new monarch. The end of the 18th and the start of the 19th centuries saw turmoil unleashed throughout Europe by the French Revolution

and Napoleonic Wars, which led to French occupation of much of the continent, including Spain.

This triggered a successful but devastating war for Spanish independence that shattered the country and created an opening for what would ultimately be the successful independence of Spain's mainland American colonies. Fragmented by the war, Spain was destabilized as different political parties representing "liberal", "reactionary" and "moderate" groups throughout the remainder of the century fought for and won short-lived control without any being sufficiently strong to bring about lasting stability. Nationalist movements emerged in the last significant remnants of the old empire (Cuba and the Philippines) which led to a brief war with the United States (1898) and the loss of the remaining old colonies at the end of the century. It was only in the constitution of 1812 that the name "Españas" for the Spanish kingdom and the use of the title of "King of the Spain" were officially adopted. The constitution of 1876 adopted for the first time the name "Españas" ("Spain") for the Spanish nation, and from then on monarchs used the title of "King of Spain".

Following a period of growing political instability in the early 20th century, in 1936 Spain was plunged into a bloody civil war. The war ended in a nationalist

dictatorship, led by Francisco Franco, which controlled the Spanish government until 1975. Spain was officially neutral during World War II (1939-1945), although many Spanish volunteers fought on both sides. The post-war decades were relatively stable (with the notable exception of an armed independence movement in the Basque Country), and the country experienced rapid economic growth in the 1960s and early 1970s. The death of Franco in 1975 resulted in the return of the Bourbon monarchy headed by Prince Juan Carlos. While tensions remain (for example, with Muslim immigrants and in the Basque region), modern Spain has seen the development of a robust, modern democracy as a constitutional monarchy with popular King Juan Carlos, one of the fastest-growing standards of living in Europe, entry into the European Community, and the 1992 Summer Olympics and 1982 World Cup. (*www.wikkipedia.org*)

### 1.3.2 HISTORICAL BACKGROUND OF THE STUDY AREA

England is a country that is part of the United Kingdom. It shares land borders with Scotland to the north and Wales to the west the Irish Sea is to the north west, the Celtic Sea to the south west, with the North Sea to the east and the English Channel to the south separating it from continental Europe. Most of England comprises the central

and southern part of the island of Great Britain in the North Atlantic. The country also includes over 100 smaller islands such as the Isles of Scilly and the Isle of Wight.

The area now called England was first inhabited by modern humans during the Upper Palaeolithic period, but it takes its name from the Angles, one of the Germanic tribes who settled during the 5th and 6th centuries. England became a unified state in AD 927, and since the Age of Discovery, which began during the 15th century, has had a significant cultural and legal impact on the wider world. The English language, the Anglican Church, and English law—the basis for the common law legal systems of many other countries around the world—developed in England, and the country's parliamentary system of government has been widely adopted by other nations. The Industrial Revolution began in 18th-century England, transforming its society into the world's first industrialized nation. England's Royal Society laid the foundations of modern experimental science.

England's terrain mostly comprises low hills and plains, especially in central and southern England. However, there are uplands in the north (for example, the mountainous Lake District, Pennines, and Yorkshire Dales) and in the south west (for example, Dartmoor and the Cotswold's). London, England's capital is the largest metropolitan area in the United Kingdom and the largest urban zone in the European

Union by most measures. [Note 3] England's population is about 51 million, around 84% of the population of the United Kingdom, and is largely concentrated in London, the South East and conurbations in the Midlands, the North West, the North East and Yorkshire, which each developed as major industrial regions during the 19th century. Meadowlands and pastures are found beyond the major cities.

The Kingdom of England—which after 1284 included Wales—was a sovereign state until 1 May 1707, when the Acts of Union put into effect the terms agreed in the Treaty of Union the previous year, resulting in a political union with the Kingdom of Scotland to create the new Kingdom of Great Britain. In 1801, Great Britain was united with the Kingdom of Ireland through another Act of Union to become the United Kingdom of Great Britain and Ireland. In 1922, the Irish Free State was established as a separate dominion, but the Royal and Parliamentary Titles Act 1927 reincorporated into the kingdom six Irish counties to officially create the current United Kingdom of Great Britain and Northern Ireland.

The name "England" is derived from the Old English name Engla land, which means "land of the Angles". The Angles were one of the Germanic tribes that settled in Great Britain during the Early Middle Ages. The Angles came from the Angeln peninsula in the Bay of Kiel area of the Baltic Sea. According to the Oxford English

Dictionary, the first known use of "England" to refer to the southern part of the island of Great Britain occurs in 897, and its modern spelling was first used in 1538.

An alternative name for England is Albion. The name Albion originally referred to the entire island of Great Britain. The earliest record of the name appears in the Aristotelian Corpus, specifically the 4th century BC De Mundo] "Beyond the Pillars of Hercules is the ocean that flows round the earth. In it are two very large islands called Britannia; these are Albion and Ierne". The word Albion (Ἀλβίων) or insula Albionum has two possible origins. It either derives from a cognate of the Latin albus meaning white, a reference to the white cliffs of Dover, the only part of Britain visible from the European Continent, or from the phrase in Massaliote Periplus, the "island of the Albiones". Albion is now applied to England in a more poetic capacity. Another romantic name for England is Loegria, related to the Welsh word for England, Lloegr, and made popular by its use in Arthurian legend.

## 1.4 STATEMENT OF THE PROBLEM

This study seeks to analyze data on Statistical analysis of goals scored in Spain league and England league between the years 2002 - 2012. A case study of Spanish Laligal and English premiership league

## 1.5 SCOPE AND COVERAGE

This project work is focused on the Statistical analysis of goals scored in Spain league England league between the years 2002 - 2012. A case study of Spanish Laligal and English premiership league

## 1.6 AIMS AND OBJECTIVES

1. To know the linear dependency of goal scored on time.
2. To know the number of goals scored in every season.
3. To determine the degree of relationship if there is any.
4. To know the highest and lowest number of goal scored in the years of study.
5. To forecast for ten years.
6. To determine whether there is any relationship between them.
7. To know if there is any difference in the mean goal scored in the two leagues.

## 1.7 RESEARCH QUESTION

The research will sought to provide answers to the following questions:

i. The research question for this project is to check if there will be any linearity between goal scored and time.

ii. Also to test for the significant of correlation of goal scored and time.

## 1.8 RESEARCH HYPOTHESIS

The following hypothesis will be tested during data analyses at 0.05 level of significant.

**Hypothesis one:**

Ho1: There is no linearity between goal scored in Spanish Laligal and Time

**Hypothesis two:**

Ho2: The correlation is not significant between goals scored in Spanish Laligal and Time

**Hypothesis three:**

Ho3: There is no linearity between goal scored in English Premiership and Time

**Hypothesis four:**

Ho4: The correlation is not significant between goals scored in English Premiership and Time

**Hypothesis five:**

Ho5: There is no significant difference in the mean goal scored in the two leagues

# CHAPTER TWO

## LITERATURE REVIEW

The recent history of European and Spanish football is marked by the occurrence of a number of phenomena that produce effects on the ability of clubs compete with each other and able to generate profits to attract new investors, to be reinvested in the purchase of new players, improve salaries and infrastructure. During this period, weave cases of emblematic clubs that have gone bankrupt (e.g. Leeds United), clubs win the national championship but show losses (e.g. Afc Ajax during 2001/02) and clubs who have dispersed their capital in the market and who have seen recently its shares to be acquired by large investors (e.g. Manchester United). In this context it is necessary understand if the football clubs need to be profitable and if money can buy sporting performance.

This empirical study aims to study whether sports managers of English clubs who participated in the Premier League over 12 seasons, conciliating get a good sporting performance with financial performance. In this study, structural equation model was

used to analyze relationship between the three constructs early mentioned. This methodology present some advantages compared to the more commonly used techniques in studies about this field, such as regression analysis. Additionally, in the level of sporting performance, we considerate sports results obtained in national and UEFA competitions. The main conclusion to be drawn from this study is that the managers of English football clubs sought to combine sporting performance with financial performance.

The high degree of correlation estimate (0.95) between the two constructs to

Corroborate the view of Vroom an (2000) that managers try to achieve simultaneously financial and sports objectives. Another interesting result is that the association Between Turnovers and major competitions (Premier League and UEFA) is stronger compared with wages and salaries. In fact, the increase in turnovers has been accompanied by a marked increase in operational expenses.

Although in recent years the revenues of clubs are increasing, the football club failed to increase proportionately the wealth (income generated) offered to shareholders. Thus, this study to corroborate the opinion of Gerrard and Dobson (2000) that the sports managers seek to achieve a minimum level of profit and maximize sporting performance. Even in situations where the club is owned by a group of investors, the

objectives do not change. The objective of profit maximization stated by economic theory for the company is replaced by a financial goal of ensuring sufficient revenues to cover operating costs and invest in the acquisition and maintenance of the best players. As regards the directionality of the relationship, the study shows that good performance in UEFA club has immediate positive Turnovers in the four major English clubs. In the case of Chelsea the causality also includes the positive contemporary effect of the performance in the Premier League in their Turnovers. For Manchester United and Liverpool we highlight the lack of causality between sporting performance with investment in new players and costs of wages and salaries. Thus, the study contradicts the results of previous study of Dobson and Goddard (1998) which reveals that financial performance has a positive effect on sporting performance.

Finally, the study reveals a moderate correlation between stock market return and financial performance and sporting performance. This is a signal that some of the observed variables are considered by investors or there other factors that explain stock returns.

## TERMS IN FOOTBALL

***18 YARD BOX:*** Another way to say PENALTY BOX

***ABANDONED MATCH:*** A match that the REFEREE stops before REGULATION TIME (and sometimes before the game starts), for example because of very bad weather, or a bad PITCH INVASION

***ABOVE THE RELEGATION ZONE:*** In a position that will mean that you don't have to worry about being RELEGATED to a lower DIVISION if you are still in that position at the end of the SEASON.

***ACCURACY:*** Putting the ball where you want to go usually used about Passes and SHOTs. Often used in contrast to power.

***AERIAL SKILLS:*** Being good at heading the ball

***AET*** A short form of AFTER EXTRA TIME

***AFTER EXTRA TIME:*** The score after 90 minutes plus an extra thirty minutes, because the score was EVEN after 90minutes

***AGAINST THE RUN OF PLAY:*** A goal scored when the other team seems to be doing much better, for example when most of the play is in the other half of the pitch

***AGENT:*** Usually used to talk about the people who represent the players, e.g. during TRANSFER negotiations

***AGGREGATE SCORE:*** Calculated by adding the two scores together when teams play two matches, one at HOME and one AWAY, for example in the SEMI FINALS

***ALL:*** Used to show that the FINAL SCORE was a draw, e.g. “Two all” for “Two two”

***ALL-SEATER:*** A stadium with no standing room on the TERRACES where everyone must sit down instead. All British grounds in the top divisions have to be all-seaters due to safety concerns

***ANTICIPATION:*** Being able to predict what will happen next, for example where a STRIKER will try to score

***AREA (THE):*** The rectangular part of the pitch near the goal within which the GOALKEEPER is allowed to use their hands to control the ball. More properly called the PENALTY AREA

***ARTIFICIAL TURF:*** TURF0020 made from man-made materials rather than the usual grass. It is easier to maintain but difficult to play on and can injure players who fall on it

***ASSISTANT REFEREE:*** The modern official name for a LINESMAN

***ASSOCIATION FOOTBALL:*** The official name of football, to contrast it with rugby football. The informal form SOCCER is a shortening of the word “association” in this expression

***ATTACKER:*** (1) Any player with the ball who is trying to move it towards the OPPOSITION's goal (2) A striker

***ATTEMPT:*** Usually means a failed SHOT AT GOAL

***AUTOMATIC BOOKING :*** A YELLOW CARD or RED CARD that the referee has to give due to the rules, with no place to make their own judgement on the matter, e.g. a LATE TACKLE from behind with STUDS UP

***AUTOMATIC PROMOTION:*** Teams who can go up to the next DIVISION without needing to go through a PLAYOFF match with other candidates for PROMOTION, usually because they finished first or second in their division at the end of the REGULAR SEASON

***AWAY FANS:*** SUPPORTERs of the team that is not playing at their own GROUND, almost always fewer in number than the HOME FANs

***BACK FOUR:*** The DEFENDERs of one team

***BACK HEEL:*** Using the back of your foot to PASS the ball, or sometimes SHOOT

***BACK OF THE NET (IN THE):*** A goal, as a ball which goes into the goal is usually trapped at the back of the net until it is picked up

***BALLOON SHOT:*** A SHOT that went far too high

***BEAT THE KEEPER:*** Score a goal

***BENCH WARMER:*** A player who is always a substitute and rarely used

***BICYCLE KICK:*** A shot in which the player flips their body over to SHOOT (or more unusually CROSS) by kicking the ball while it is over their own head

***BID FOR A PLAYER:*** A team saying how much they would pay the team that a player has a contract with if that player was to TRANSFER to their club

***BOTTOM OF THE TABLE:*** The 12th team in a twelve team DIVISION, likely to suffer an automatic DEMOTION to a lower division if they don't win soon

***CAPTAIN:*** A player who is given responsibility for organizing the other players on the pitch

***CAPTAIN'S ARMBAND:*** A piece of fabric worn around the upper arm that shows who the CAPTAIN of each team is. If the captain is SUBSTITUTED, this is passed to another player.

***CARDS:*** YELLOW CARDs and RED CARDs given for breaking the rules of the game. Often used to show how badly one or both sides broke the rules during the match, e.g. "Last night's match produced a record of 12 cards, including 4 red ones, for the two sides"

***CENTRAL DEFENDERS:*** The one or (usually) two DEFENDERS who plays near the middle of the PITCH rather than on the WINGS

***CENTRAL MIDFIELD:*** MIDFIELDERS who play near the centre circle rather than on the WING

***CENTRE CIRCLE:*** The circle in the centre of the PITCH, with a 10 yard radius centered on the **CENTRE SPOT:** The KICK OFF takes part at the centre of this, with players from the opposing side staying outside this circle until the whistle for kick off is blown.

***CENTRE FORWARD:*** The FORWARD who plays nearest to the goal and so is often the TARGET MAN

***CENTRE HALF:*** Another way to say CENTRAL DEFENDER

***CENTRE SPOT:*** The point right in the middle of the pitch where you KICK OFF from

***CLEAN TACKLE:*** Taking the ball from someone without it being a FOUL

***CLEARANCE*** Getting the ball away from your own goal, e.g. by kicking it far away towards the other goal

***CLIMB:*** Jump to HEAD the ball

***CROSSBAR:*** The piece of wood that goes horizontally across the top of the goal, supported by the two POSTs. “The ball hit the crossbar” usually means that it was nearly a goal.

***CUP CHAMPION:*** The opposite of a LEAGUE CHAMPION, a team that has won a CUP COMPETITION

***DEFENDER:*** A person whose main role is to stop the opposition's attacking players from scoring and getting the ball near their goal

***DEFENSIVE MIDFIELDER:*** A midfielder whose main role is to PLAY DEEP or SHIELDS the defenders, i.e. whose main duties are defensive

***DEMOTION:*** The same as RELEGATION. Being forced to go down to a lower DIVISION because you finished last at the end of season

***DERBY:*** A match between sides who play in grounds that is close to each other, for example in the same or neighboring towns

***DISALLOWED GOAL:*** The time when the ball goes into the net but the referee says it was not a goal, for example because a player was OFFSIDE

***DRESSING ROOM:*** The place where the players get changed before and after a match. Also called CHANGING ROOMS and LOCKER ROOM

***DRIBBLE:*** Controlling the ball with your feet as you run

***DROP BACK*** Players playing for a while closer to their own goal than usual, e.g. to defend against an attacking move

***EQUALISER:*** A goal which makes the scores EVEN, for example going from 2-1 to 2-2

***FAN:*** A supporter of a particular club. Although fan comes from the word FANATIC, their modern meanings are very different.

***FIRST HALF:*** The first 45 minutes of match (plus INJURY TIME), followed by a break

***FORMATION:*** How the players are arranged on the pitch, e.g. 4-4-2

***FORWARD:*** A more official way of saying STRIKER

***FROM 12 YARDS:*** From the PENALTY SPOT (12 yards is the distance from the goal to the penalty spot).

***GOAL:*** (1) The rectangular opening consisting of a horizontal CROSSBAR supported by two bars, usually with a net behind. (2) Kicking the ball into the goal of the other team

***HAND BALL:*** A foul in which the ball hits the hand or lower parts of the arm.

***HAT TRICK:*** Three goals by the same player in the same match

***HEADER:*** Using your head to control the ball, e.g. in trying to score a goal. Typical mistake heading shoot

***LINESMAN:*** The traditional name for the two ASSISTANT REFEREES who run up and down outside the SIDELINES, rather than around on the pitch like the REFEREES. Linesmen often decide if a player is OFFSIDE or if the ball has gone OUT OF PLAY

***MAN TO MAN MARKING:*** A system in which each defender is given one attacking player from the other team that they have to stay close to. The opposite of ZONAL MARKING

***MANAGER***: Traditionally, someone who is in charge of picking a team, deciding on tactics, and buying and selling players, one level higher in management than a COACH.

***MARKING:*** Staying close to an attacking player to make sure they can't get the ball or can't PASS or SHOOT if they have it.

# CHAPTER THREE

## METHODOLOGY OF THE STUDY

The tools employed by researcher in the project include following:

The charts and graphs are used for the presentation of the data, while correlation analysis, regression analysis, test of hypothesis and trend estimation are for the purpose of inferred statistics which is the beauty of any statistical enquiry and Analysis of Variance (ANOVA) table. Also Student T distribution will be use to determine the mean difference.

### 3.1 CHARTS AND GRAPHS

Charts and graphs are diagrammatic representation of data in order to show relationship or compare magnitude of two or more variables. The charts and graphs used in this project are simple bar chart and line graph.

#### 3.1.1 SIMPLE BAR CHART

In this method, bars are drawn proportionally to the sizes of the items of interest. It is exposing because one can easily see the size of each of the item at a glance.

#### 3.1.1 LINE GRAPH

A line chart or line graph is a graphical device that displays quantitative information or illustrates relationships between two changing variables with a line or curve that connects a series of successive data points.

## 3.2 CORRELATION ANALYSIS

Correlation can be defined as the degree of association between two or more variables whereby the movement in one of other variable is associated with the movement of the other variable within the same direction or opposite direction. It is also a relationship between two variables such that the value of one is affected by the value of the other.

When only two variables are involved, it is simple correlation, but when more than variable are involved, it is multiple correlations.

In this project, simple correlation was used was used because it deals with two variables X and Y.

Correlation between variables X and Y is perfect if all the points of the scatter diagram lie on a straight line, i.e. when $r = \pm 1$.

Correlation between two variables X and Y is positive if Y increases as X increases or if the regression line slopes upward from left to right i.e. , r=+1. Correlation between two variables X and Y is negative or inverse. If Y decreases as X increases or Y increases as

X decreases as correlation between two variables, X and Y if there is no define pattern in the direction of the variable X and Y, i.e. r = 0.

Correlation coefficient shows that extent or magnitude of relation that exists and its value ranges between -1 and +1, i.e. $-1 \leq r \leq +1$.

The formula for calculating 'r' can be expressed by the formula.$r = \frac{n\sum XY-(\sum X(\sum Y))}{\sqrt{n\sum X^2-(\sum X)^2}(n\sum Y^2-(\sum Y))^2}$

Where r= the correlation coefficient

N= number of observation

X= variable

Y= variable

**INTERPRETATION OF CORRELATION COEFFICIENT**

i. When r = 1, there is a perfect positive correlation

ii. When r = -1, there is a perfect negative correlation

iii. When 0.5<r<1, there is a strong positive correlation

iv. When -1<r<-0.5, there is a strong negative correlation

v. When 0<r>0.5, there is a weak positive correlation

vi. When -0.5<r<0, there is a weak negative correlation

vii. When r = 0, there is no linear correlation

### 3.2.1 COEFFICIENT OF DETERMINANTION ($r^2$)

This shows how much of the dependent variable is explained by the independent variable. It is an increasing function whose values lie between 0 and 1, i.e. $0 \leq R^2 \leq 1$.

## REGRESSION ANALYSIS

Regression is a statistical method of forming a model of the relationship between variables in order to enable prediction or estimate to be made.

The regression equation is usually of the form Y = a +bx. By partial derivatives, the normal equations, from which the values of the parameters $\partial$ and β can be estimated, are derived thus:

Differentiating $\sum e^2$with respect to "a" keeping b constant $\partial \sum e^2 / \partial\, a = 2 \sum [y - a - \text{bx}] = 0$

$$-\sum (Y - a - \text{bx}) = 0$$

$$\left(-\sum y + \text{na} + b \sum x\right) = 0.$$

$$-\sum \text{yna} + b \sum x \ldots \text{equ}(i)$$

Differentiating $\sum e^2$with respect to "b" keeping a constant

$$\partial \sum e^2/\partial\, a = -x2 \sum (y - a - \text{bx}) = 0$$

$$= -2 \sum (y - a - \text{bx})$$

$$\sum (y - a - \text{bx})x = 0$$

$$\sum \text{xy} = a \sum x + b \sum x2 \text{...equ(ii)}$$

by combining equation (i) and (ii)

$$\sum y = \text{na} + b \sum x \text{.....}(i)$$

$$\sum \text{xy} = a \sum x + b \sum x^2 \text{.....(ii)}$$

Multiply equation (i) by x and equation (ii) by n we have $\sum x \sum y = \text{na} \sum x + b(\sum x)^2$.....(iii)

$$n \sum \text{xy} = \text{na} \sum y + \text{nb} \sum x^2 \text{.....(iv)}$$

Subtracting equation (iii) from (iv)

$$n \sum \text{xy} - \sum x \sum y = \text{na} \sum x^2 - b \left( \sum x \right)^2$$

$$n \sum \text{xy} - \sum x \sum y = b \left( n \sum x^2 - b \left( \sum x \right)^2 \right)$$

$$b = \frac{n \sum \text{xy} - \sum x (\sum y)}{n \sum x^2 - (\sum x)^2}$$

By replacing b in the equation (i)

$$\sum y = \frac{\text{na} + n \sum \text{xy} - \sum x (\sum y) \sum x}{n \sum x^2 - (\sum x)^2}$$

By dividing by n

$$\frac{\sum y}{n} = \frac{\text{na} + n \sum \text{xy} - \sum x (\sum y) \sum x.}{n(n \sum x^2 - (\sum x)2)}$$

$$\frac{\sum y}{n} = \frac{a + \sum \text{xy}(n\sum \text{xy} - \sum x(\sum y))}{n(n\sum x^2 - (\sum x)2)}$$

$$a = \frac{\sum y}{n} - \frac{\sum x(n\sum \text{xy} - \sum x(\sum y))}{n(n\sum x^2 - (\sum x)^2)}$$

$$a = \overline{y} - \overline{x}b$$

Where 'a' is the intercept on the y-axis but b is the slope or gradient of the regression line. The "b" is also called the coefficient of regression. This is the error term.

The regression line is sometimes called least square regression line.

### 3.3.1 SCATTER DIAGRAM

The point that indicates the co-ordinates of the variable under consideration are shown on a graph called scatter diagram. Each variable x and y in the diagram is subject to considerable variability so that we obtain a scattered point.

### 3.3.2 REGRESSION EQUATION

This is the equation that gives linear relationship between two variables given a value of variable. We can estimate for the corresponding value of the other variable using equation given by Y = a + bx

### 3.3.3 REGRESSION LINE

Line of best fit is the line that best describes the general direction of the scatter diagram. The line in which the independent of individual judgment will have to be drawn mathematically is the regression line.

The equation of the line is $Y = a + bx$.

To describe the line parameter a and b would be estimated. The parameter a and b could be obtained using the equation which can in turn be obtained by partial derivation. This procedure is the least square method.

## 3.4 STATISTICAL HYPOTHESIS

In the process of decision making, it is necessary to make assumptions or guess about the population involved.

This assumption may or may not be true. It is claim or assertions about a particular phenomenon of interest.

**Null Hypothesis:** States that there is no difference between the hypothesis value and the value obtained. It is that hypothesis that I want to accept or reject and it is denoted by Ho.

**Alternative Hypothesis:** States that there is difference between the hypothetical value and the population value, which is denoted H1.

**Level of significance**: The maximum probability with which one will be willing to risk type 1 error is called the level of significance.

**Type 1 Error:** Rejection of the null hypothesis when it is true is called a type 1 error. The probability of committing a type 1 error is denoted by α, i.e., if hypothesis is rejected when it should be accepted.

**Type II Error:** The error is committed when the null hypothesis is accepted when it ought to have been rejected.

**Test statistic:** this is computed from the sample data in order to decide whether to reject or accept a null hypothesis.

**Critical region:** all possible set of values of the test statistic that lie to one side of the critical value constitute the rejection region, while those values that lie in the other side of the critical value constitute the acceptance region.

**Degree of freedom:** it is the sample size minus the number k of population parameter estimate the sample i.e., df = n-k.

## Test of Hypothesis

Hypothesis is not based only on scientific research, but too many business situations where simple data are used to evaluate the effect of change to support or proposed on a policy. These are the procedure for testing of hypothesis. Firstly, state the null hypothesis, Ho, there is no difference between the procedures and the alternative hypothesis, H1 there is difference between the procedures. Formulate a statistical hypothesis for the rejecting or nullifying purposes.

The next is to define the test statistic, which helps in decision making. This could be student's t test of Z test depending on the size of the sample. The level of significance of the test denoted by $\alpha$ has to be stated before the conclusion. In this project all the test was based on $\alpha = 0.05$ (5% level of significance). In scientific research, I developed theories and then conduct experiment.

## TESTING THE HYPOTHESIS ABOUT LINEARITY

The population coefficient of regression $\beta$ is estimates by the sample coefficient of regression.

The usual significant test is where b is equal to some specified value $\beta$.

The hypothesis is set as follows:

$H_o$: $\beta = 0$ (i.e. x and y are not linearly relation)

$H_1$: $\beta \neq 0$ (i.e. x and y are linearly related)

Calculation:

$$SST = total \sum of\ square = \sum y^2 -$$

$$SSR = regressionsunofsquare = b\left\{\sum xy - \frac{\sum x \sum y}{n}\right\}$$

SSE = Sum of square error = SST - SSR

From the notations and formula analysis of variance table (ANOVA TABLE) can be constructed as below:

| **SOURCE OF VARIATION** | **SUM OF SQUARE** | **DEGREE OF FREEDOM** | **MEAN SQUARE** | **F-RATIO** |
|---|---|---|---|---|
| Regression | $b\left\{\sum xy - \frac{\sum x \sum y}{n}\right\}$ | 1 | $\text{MSR} = \frac{\text{SSR}}{1}$ | $f = \frac{\text{MSR}}{\text{MSE}}$ |
| Error | $SST - SSR$ | n-2 | $\text{MSE} = \frac{\text{SSR}}{N-2}$ | |

| | | | | |
|---|---|---|---|---|
| Total | $\sum y^2 -$ | n-1 | | |

**Conclusion:** If the calculated F value is greater than the tabulated F value at the given or particular level of significance, then Ho cannot be accepted. Hence, we accept H1 and vice versa.

**STUDENT *T-Test***

This is normally used to compare the *Means* of two groups of data; i.e. the data being compared should be quantitative. These two groups of data may be for two independent samples or may be for the same Sample with the data collected at two different periods {i.e. paired samples}. If, based on the observed p-value, it is decided that the two groups are Different, then, one should be able to state which group has the larger *Mean*.

The test is used to compare two sample distributions, which should be approximately normal, with data collected on either interval or ration scale of measurement. For a two sample test, the student t test is given by:

$$t = X - \frac{Y/}{\sqrt{\frac{\sigma x^2}{Nx} + \frac{\sigma y^2}{Ny}}}$$

**Where** **t = student t test**

**X = sample No. 1**

**Y = sample No. 2**

$\sigma x^2$ **= a square of the standard deviation of X**

$\sigma y^2$ = **square of the standard deviation of Y**

$Nx$ = **sample size of X**

$Ny$ = **sample size of Y**

The higher the value of t, the higher the discrepancy between the means of the sample and the population.

In the students't-test, a comparison of two continuous variables is made, using their group means as the basis for the comparison. Thus the goal is to find out whether or not a significant difference between two samples truly implies a difference between the means of the parent populations.

# CHAPTER FOUR

## DATA PRESENTATION AND ANALYSIS

**TABLE 1**

**SOURCE:** **http://www.betstudy.com/soccer-stats/c/spain/primera-division, http://www.betstudy.com/soccer-stats/c/england/premier-league**

| YEARS | GOAL SCORED IN LALIGAL | GOAL SCORED IN ENGLISH PREMIER LEAGUE |
|---|---|---|
| 2002/2003 | 1016 | 1000 |
| 2003/2004 | 1015 | 1012 |
| 2004/2005 | 980 | 975 |
| 2005/2006 | 936 | 944 |
| 2006/2007 | 942 | 931 |
| 2007/2008 | 1021 | 1002 |
| 2008/2009 | 1101 | 942 |
| 2009/2010 | 1031 | 1053 |
| 2010/2011 | 1042 | 1063 |
| 2011/2012 | 1050 | 1066 |

## DATA ANALYSIS AND INTERPRETATION

## REGRESSION ANALYSIS OF GOAL SCORED AND TIME IN SPANISH LALIGAL

| YEAR | X | Y | $X^2$ | $Y^2$ | XY |
|---|---|---|---|---|---|
| 2002/2003 | 1 | 1016 | 1 | 1032256 | 1016 |
| 2003/2004 | 2 | 1015 | 4 | 1030225 | 2030 |
| 2004/2005 | 3 | 980 | 9 | 960400 | 2940 |
| 2005/2006 | 4 | 936 | 16 | 876096 | 3744 |
| 2006/2007 | 5 | 942 | 25 | 887364 | 4710 |
| 2007/2008 | 6 | 1021 | 36 | 1042441 | 6126 |
| 2008/2009 | 7 | 1101 | 49 | 1212201 | 7707 |
| 2009/2010 | 8 | 1031 | 64 | 1062961 | 8248 |
| 2010/2011 | 9 | 1042 | 81 | 1085764 | 9378 |
| 2011/2012 | 10 | 1050 | 100 | 1102500 | 10500 |
| **Total** | **55** | **10134** | **385** | **10292208** | **56399** |

N =10 $\sum xy = 56399$ $\sum x = 55$ $\sum y = 10134$ $\sum x^2 = 385$

$$b = \frac{n\sum \mathrm{XY} - (\sum X)(\sum Y)}{n\sum X^2 - (\sum X)^2}$$

$$b = \frac{10(56399) - (55)(10134)}{10(385) - (55)^2}$$

$b = 8.024$

Where a $= \overline{y}$- b$x$ $\quad \overline{y} = \frac{\sum y}{n}$ and $\quad \bar{X} = \frac{\sum Xi}{n}$

$\overline{y} = \frac{10134}{10} = 1013.4$ and $\bar{X} = \frac{55}{10} = 5.5$

a $= \overline{y}$- b$x$

a = 1013.4 – (8.024) (5.5)

a = 969.268

$$Y = a + bx$$

$$Y = 969.268 + 8.024x$$

**INTERPRETATION**

The regression equations obtained above is;

$Y = 969.268 + 8.024X$ Since the value of "b" is positive, for every increase in x which is Time, there will be a corresponding increase in the number of **GOAL SCORED** in Spanish Laligal. While in the absent of time i.e. when x=0, y= 969

**TEST FOR LINEARITY BETWEEN TIME AND GOAL SCORED**

$H_o$: a = b (there is no linearity between time and goal scored)

$H_1$: a ≠ b (there is linearity between time and goal scored)

**Critical region**

$F\propto,(v1,v2)$

$F0.05,(1,8) = 5.32$

**Decision:** Accept Ho if $F_{cal} < F_{tab}$ or otherwise reject

**Test statistic**

$SSR = b\frac{\sum x \sum y}{n})$

$SSR = 8.024(\mathbf{56399} - \frac{557370)}{10}$

$= 8.024\ (\mathbf{56399} - 55737)$

$= 8.024(662)$

$SSR = 5311.888$

$SST = \sum y^2 -$

$= \mathbf{10292208} - \frac{(10134)^2}{10}$

$=\mathbf{10292208} - 10269795.6$

$SST = 22412.4$

SSE = SST – SSR

SSE = 22412.4 – 5311.888

SSE = 17100.512

## ANOVA TABLE

| SOURCE OF VARIATION | SUM OF SQUARES | DEGREE OF FREEDOM | MEAN SQUARE | F-RATIO |
|---|---|---|---|---|
| TREATMENT | 5311.888 | 2 – 1 = 1 | 5311.888 | 2.485 |
| ERROR | 17100.5 | 10 – 2 = 8 | 2137.5625 | |
| TOTAL | 22412.4 | 10 – 1 = 9 | | |

**Critical region**

$$\alpha = 0.05 confidence interval$$

$$1 - \propto = significance level$$

$$1 - 0.05 = 0.95 significance level$$

$F0.05, (1,8) = 5.32$

$F calculated = 2.485$

**Decision rule:** If $F_{calculated} < F_{tabulated}$, we accept $H_o$ otherwise reject

**Conclusion:** - since $F_{calculated\ (2.485)} < F_{tabulated\ (5.32)}$; we therefore accept $H_o$ and conclude that there is no linearity between time and goal scored. Therefore we cannot forecast with the linear modeled obtained.

## CORRELATION ANALYSIS OF GOAL SCORED AND TIME

$$r = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{[n\sum X^2 - (\sum X)^2][n\sum Y^2 - (\sum Y)^2]}}$$

$$r = \frac{10(56399) - (55)(10134)}{\sqrt{[10(385) - (55)^2][10(10292208) - (10134)^2]}}$$

$\sigma$= 0.4868

**INTERPRETATION:**

The value above r=0.4868 implies that there is weak positive correlation between **TIME** and **GOAL SCORED** in Spanish Laligal. The degree of association is about 48.68%. This means that Increase in **time** will bring about increase in **goal scored**.

**COEFFICIENT OF DETERMINATION ($r^2$)**

Where $r^2 = (0.4868)^2 * 100$

$r^2 = 0.2370$

**INTERPRETATION:**

The co-efficient of determination $r^2$ which is 0.4868 and about 24% of **GOAL SCORED** in Spanish Laligal can be explained by **TIME** leaving about 76% to be explained by other factor.

**SIGNIFICANT TEST FOR CORRELATION COEFFICIENT**

$\sigma$=0.4868 from x and y, n = 10,$\propto$= 0.05

## Hypothesis

Ho:$\rho = 0(thecorrleationisnotsignificant)$

H1: $\rho \neq 0(thecorrleationissignificant)$

**Critical Region:**

$t\frac{\propto}{2}, (n-2)d.f$

$t_{tab}$ = 2.306

**Test statistic:**

$$t = \frac{r\sqrt{n-2}}{\sqrt{[1-r^2]}}$$

$$t = \frac{0.4868\sqrt{10-2}}{\sqrt{[1-(0.4868)^2]}}$$

$$t = \frac{0.4868\sqrt{8}}{\sqrt{[1-0.2370]}}$$

$t_{cal}$ =1.5763

**Decision:** If $t_{calculated} < t_{tabulated}$ we accept Ho or otherwise

**Conclusion:** since $t_{calculated}$ (1.5763) < $t_{tabulated}$ *(2.306)* we can then accept Ho and conclude that the association between time and goal scored is not significant.

## REGRESSION ANALYSIS OF GOAL SCORED AND TIME IN ENGLISH PREMIERSHIP

| YEAR | X | Y | $X^2$ | $Y^2$ | XY |
|---|---|---|---|---|---|
| 2002/2003 | 1 | 1000 | 1 | 1000000 | 1000 |
| 2003/2004 | 2 | 1012 | 4 | 1024144 | 2024 |
| 2004/2005 | 3 | 975 | 9 | 950625 | 2925 |

| | | | | | |
|---|---|---|---|---|---|
| 2005/2006 | 4 | 944 | 16 | 891136 | 3776 |
| 2006/2007 | 5 | 931 | 25 | 866761 | 4655 |
| 2007/2008 | 6 | 1002 | 36 | 1004004 | 6012 |
| 2008/2009 | 7 | 942 | 49 | 887364 | 6594 |
| 2009/2010 | 8 | 1053 | 64 | 1108809 | 8424 |
| 2010/2011 | 9 | 1063 | 81 | 1129969 | 9567 |
| 2011/2012 | 10 | 1066 | 100 | 1136356 | 10660 |
| **Total** | **55** | **9988** | **385** | **9999168** | **55637** |

N =10 $\sum xy = 55637$ $\sum x = 55$ $\sum y = 9988$ $\sum x^2 = 385$

$$b = \frac{n\sum \mathrm{XY} - (\sum X)(\sum Y)}{n\sum X^2 - (\sum X)^2}$$

$$b = \frac{10(55637) - (55)(9988)}{10(385) - (55)^2}$$

*b = 8.5212*

Where a =$\overline{y}$- b$x$ $\overline{y} = \frac{\Sigma y}{n}$ and $\bar{X} = \frac{\Sigma \mathrm{Xi}}{n}$

$\overline{y} = \frac{9988}{10} = 998.8$ and $\bar{X} = \frac{55}{10} = 5.5$

a =$\overline{y}$- b$x$

a = 998.8 – (8.5212) (5.5)

a = 951.93

$$Y = a + \text{bx}$$

$$Y = 951.93 + 8.5212x$$

**INTERPRETATION**

The regression equations obtained above is;

$Y = 951.93 + 8.5212X$Since the value of "b" is positive, for every increase in x which is Time, there will be a corresponding increase in the number of **GOAL SCORED** in English Premiership. While in the absent of time i.e. when x=0, y= 952

**TEST FOR LINEARITY BETWEEN TIME AND GOAL SCORED**

$H_o$: a = b (there is no linearity between time and goal scored)

$H_1$: a ≠ b (there is linearity between time and goal scored)

**Critical region**

$F\propto, (v1, v2)$

$F0.05, (1,8) = 5.32$

**Decision:** Accept Ho if $F_{cal} < F_{tab}$ or otherwise reject

**Test statistic**

SSR = b$\frac{\sum x \sum y}{n}$ )

SSR = 8.5212(55637 –54934)

10

= 8.5212 (55637 – 54934)

= 8.5212(703)

SSR = 5990.4036

SST = $\sum y^2$ –

= 9999168 – $\frac{(9988)^2}{10}$

=9999168 – 9976014.4

SST = 23153.6

SSE = SST – SSR

SSE 23153.6 – 5990.4036

SSE = 17163.1964

## ANOVA TABLE

| SOURCE OF VARIATION | SUM OF SQUARES | DEGREE OF FREEDOM | MEAN SQUARE | F-RATIO |
|---|---|---|---|---|
| TREATMENT | 5990.4036 | 2 – 1 = 1 | 5990.4036 | 2.7922 |
| ERROR | 17163.1964 | 10 – 2 = 8 | 2145.340 | |
| TOTAL | 23153.6 | 10 – 1 = 9 | | |

**Critical region**

$$\alpha = 0.05 confidenceinterval$$

$$1 - \propto = significancelevel$$

$$1 - 0.05 = 0.95 significancelevel$$

$F0.05, (1,8) = 5.32$

$F calculated = 2.7922$

**Decision rule:** If $F_{calculated} < F_{tabulated}$, we accept $H_o$ otherwise reject

**Conclusion:** - since $F_{calculated\ (2.7922)} < F_{tabulated\ (5.32)}$; we therefore accept $H_o$ and conclude that there is no linearity between time and goal scored in English premier league. Therefore we cannot forecast with the linear model obtained.

**CORRELATION ANALYSIS OF GOAL SCORED AND TIME**

$$r = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{[n\sum X^2 - (\sum X)^2][n\sum Y^2 - (\sum Y)^2]}}$$

$$r = \frac{10(55637) - (55)(9988)}{\sqrt{[10(385) - (55)^2][10(9999168) - (9988)^2]}}$$

$\sigma$= 0.5087

**INTERPRETATION:**

The value above r=0.5087 implies that there is strong positive correlation between **TIME** and **GOAL SCORED** in English Premiership. The degree of association is about 50.87%. This means that Increase in **time** will bring about increase in **goal scored** since they travelled in the same pattern.

**COEFFICIENT OF DETERMINATION ($r^2$)**

Where $r^2 = (0.5087)^2 * 100$

$r^2 = 0.2588$

**INTERPRETATION:**

The co-efficient of determination $r^2$ which is 0.2588 and about 26% of **GOAL SCORED** can be explained by **TIME** leaving about 74% to be explained by other factor.

**SIGNIFICANT TEST FOR CORRELATION COEFFICIENT**

$\sigma$=0.5087 from x and y, n = 10,$\propto$= 0.05

**Hypothesis**

Ho:$\rho = 0(thecorrleationisnotsignificant)$

H1: $\rho \neq 0(thecorrleationissignificant)$

**Critical Region:**

$t\frac{\propto}{2}, (n-2)d.f$

$t_{tab} = 2.306$

**Test statistic:**

$$t = \frac{r\sqrt{n-2}}{\sqrt{[1-r^2]}}$$

$$t = \frac{0.5087\sqrt{10-2}}{\sqrt{[1-(0.5087)^2]}}$$

$$t = \frac{0.5087\sqrt{8}}{\sqrt{[1-0.2588]}}$$

$t_{cal} = 1.6712$

**Decision:** If $t_{calculated} < t_{tabulated}$ we accept Ho or otherwise

**Conclusion:** since $t_{calculated}$ (1.6712) < $t_{tabulated}$ *(2.306)* we can then accept Ho and conclude that the association between time and goal scored is not significant.

# STUDENT T ANALYSIS ON DIFFERENCE OF TWO ME

$$\sigma B = \sqrt{\frac{23153.6}{9}}$$

$$\sigma B = \sqrt{2572.622222}$$

$$\sigma B = 50.721$$

$$\sigma A = \sqrt{\sum}$$

$$\sigma A = \sqrt{\frac{22412.4}{10-1}}$$

$$\sigma A = \sqrt{\frac{22412.4}{9}}$$

$$\sigma A = \sqrt{2490.26667}$$

$$\sigma A = 49.903$$

$T_{cal} = 1013.4 - \frac{998.8/}{\sqrt{\frac{2490.2667}{10}+\frac{2572.6222}{10}}}$

$=\frac{14.6}{\sqrt{249.0267+257.2622}}$

$$\frac{14.6}{\sqrt{506.2889}}$$

$$\frac{14.6}{22.5009}$$

$T_{cal} = 0.6489$

**Hypothesis;**

$H_o$: (there is no significance difference in the number of goal scored in the two leagues)

$H_1$: (there is.)

**Degree of freedom**

NA + NB -2 at 0.05

(10 + 10) – 2 at 0.05

20 – 2 (0.05)

18(0.05)

$T_{tab}$ = 2.101

**CONCLUTION:**

Since the value of Tcal=0.65 < Ttab=2.101, we can therefore accept Ho and conclude that there is no significant difference in the number of goal scored in the two leagues.

# SPSS ANALYSIS

## CO-EFFICIENT OF DETERMINATION

**Model Summary**

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
|---|---|---|---|---|
| 1 | .487[a] | .237 | .142 | 46.234 |

a. Predictors: (Constant), Time

**INTERPRETATION:** The co-efficient of determination r2 which is 0.237 and about 24% of goal scored in Spanish Laligal can be explained by Time leaving about 76% to be explained by other factor.

## TEST OF LINEARITY USING ANOVA

**ANOVA**

| Model | | Sum of Squares | Df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| 1 | Regression | 5312.048 | 1 | 5312.048 | 2.485 | .154[b] |
| | Residual | 17100.352 | 8 | 2137.544 | | |
| | Total | 22412.400 | 9 | | | |

a. Dependent Variable: Goal scored

b. Predictors: (Constant), Time

INTERPRETATION: sig-value = 0.154 > 0.05 shows that the alternative hypothesis is rejected meaning that there is no linearity between Time and Goal Scored i.e. we cannot forecast with the linear model obtained.

Coefficients

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
|---|---|---|---|---|---|---|
| | | B | Std. Error | Beta | | |
| 1 | (Constant) | 969.267 | 31.584 | | 30.689 | .000 |
| | Time | 8.024 | 5.090 | .487 | 1.576 | .154 |

a. Dependent Variable: Goal scored

INTERPRETATION: The regression equation obtained from the table above is;

$Y = 969.268 + 8.024X$Since the value of "b" is positive, for every increase in x which is Time, there will be a corresponding increase in the number of **GOAL SCORED in Spanish Laligal**. While in the absent of time i.e. when x=0, y= 969

## CORRELATION ANALYSIS OF GOAL SCORED IN SPANISH LALIGAL AND TIME

**Correlations**

| | | year | goal scored |
|---|---|---|---|
| Year | Pearson Correlation | 1 | .487 |
| | Sig. (2-tailed) | | .154 |
| | N | 10 | 10 |
| goal scored | Pearson Correlation | .487 | 1 |
| | Sig. (2-tailed) | .154 | |
| | N | 10 | 10 |

## INTERPRETATION:

The value above r=0.487 implies that there is weak positive correlation between **TIME** and **GOAL SCORED in Spanish Laligal**. The degree of association is about 48.68%. This means that Increase in **time** will bring about increase in **goal scored**.

## CO-EFFICIENT OF DETERMINATION

**Model Summary**

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
|---|---|---|---|---|

| | | | | |
|---|---|---|---|---|
| 1 | .509[a] | .259 | .166 | 46.318 |

a. Predictors: (Constant), TIME

**INTERPRETATION:** The co-efficient of determination r2 which is 0.259 and about 26% of goal scored in English Premiership can be explained by Time leaving about 74% to be explained by other factor.

## TEST OF LINEARITY USING ANOVA

**ANOVA[b]**

| Model | | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| 1 | Regression | 5990.412 | 1 | 5990.412 | 2.792 | .133[a] |
| | Residual | 17163.188 | 8 | 2145.398 | | |
| | Total | 23153.600 | 9 | | | |

a. Predictors: (Constant), TIME

b. Dependent Variable: GOAL SCORED

**INTERPRETATION**: sig-value = 0.133 > 0.05 shows that the alternative hypothesis is rejected meaning that there is no linearity between Time and Goal Scored in English Premiership i.e. we cannot forecast with the linear model obtained.

**Coefficients[a]**

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
|---|---|---|---|---|---|---|
| | | B | Std. Error | Beta | | |
| 1 | (Constant) | 951.933 | 31.642 | | 30.085 | .000 |

| | | | | | |
|---|---|---|---|---|---|
| TIME | 8.521 | 5.099 | .509 | 1.671 | .133 |

a. Dependent Variable: GOAL SCORED

**INTERPRETATION:** The regression equation obtained from the table above is;

$Y = 951.933 + 8.521X$ Since the value of "b" is positive, for every increase in x which is Time, there will be a corresponding increase in the number of **GOAL SCORED in English Premiership**. While in the absent of time i.e. when x=0, y= 952

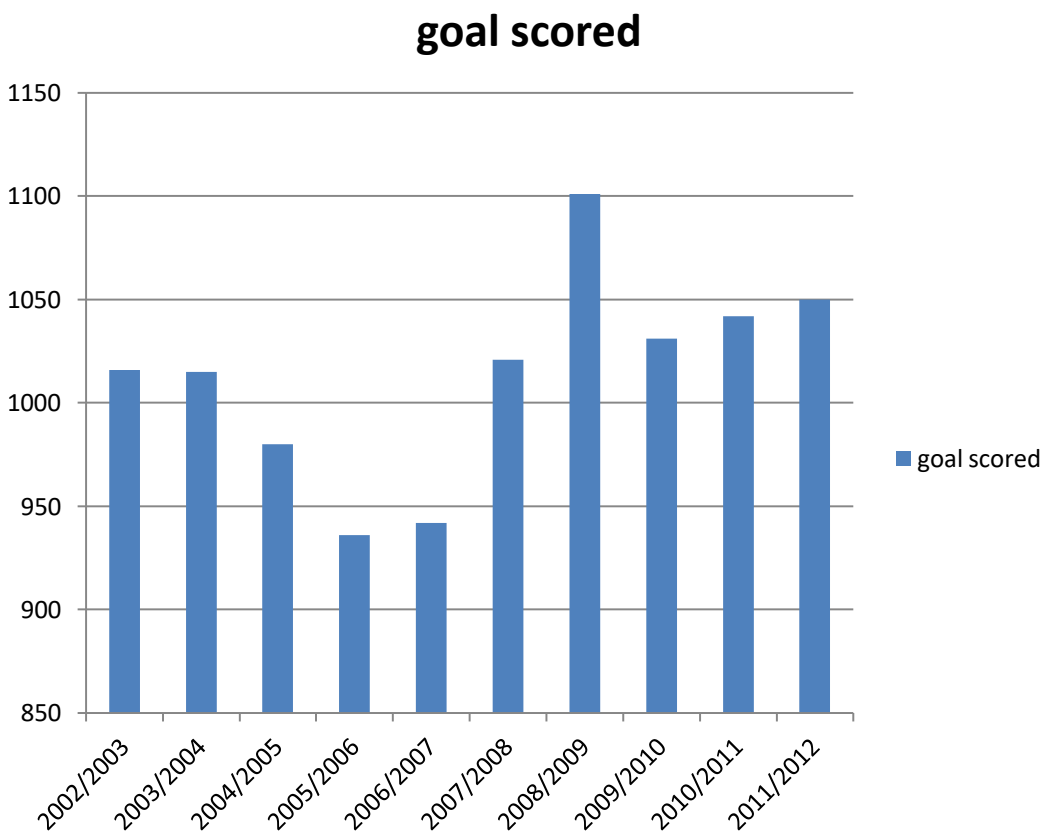
## CORRELATION ANALYSIS OF GOAL SCORED AND TIME

**Correlations**

| | | TIME | GOAL SCORED |
|---|---|---|---|
| TIME | Pearson Correlation | 1 | .509 |
| | Sig. (2-tailed) | | .133 |
| | N | 10 | 10 |
| GOAL SCORED | Pearson Correlation | .509 | 1 |
| | Sig. (2-tailed) | .133 | |
| | N | 10 | 10 |

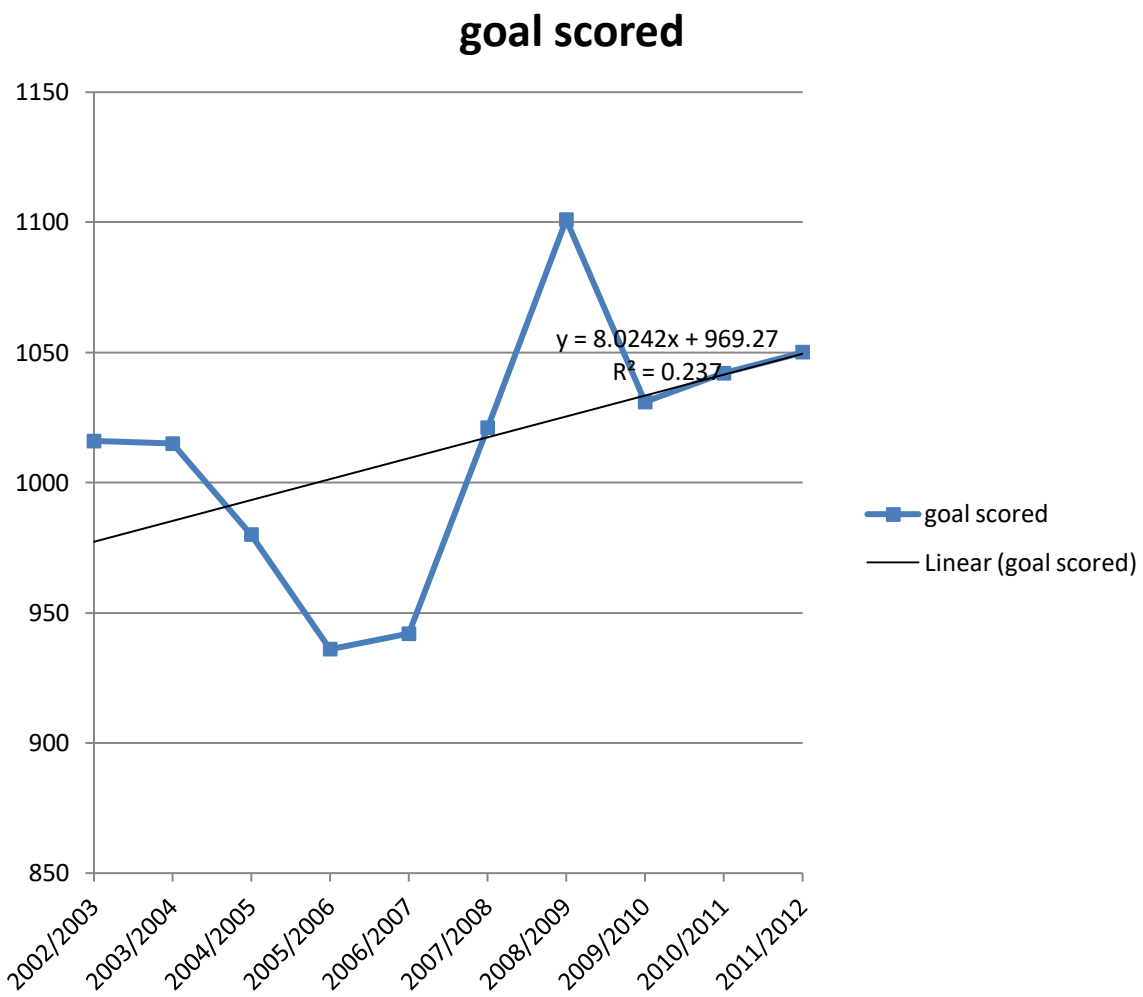
**INTERPRETATION:** The value above r=0.509 implies that there is strong positive correlation between **TIME** and **GOAL SCORED in English Premiership**. The degree of association is about 50.9%. This means that Increase in **time** will bring about increase in **goal scored**.

**SIMPLE BAR CHART SHOWING GOAL SCORED AGAINST TIME IN LALIGAL**



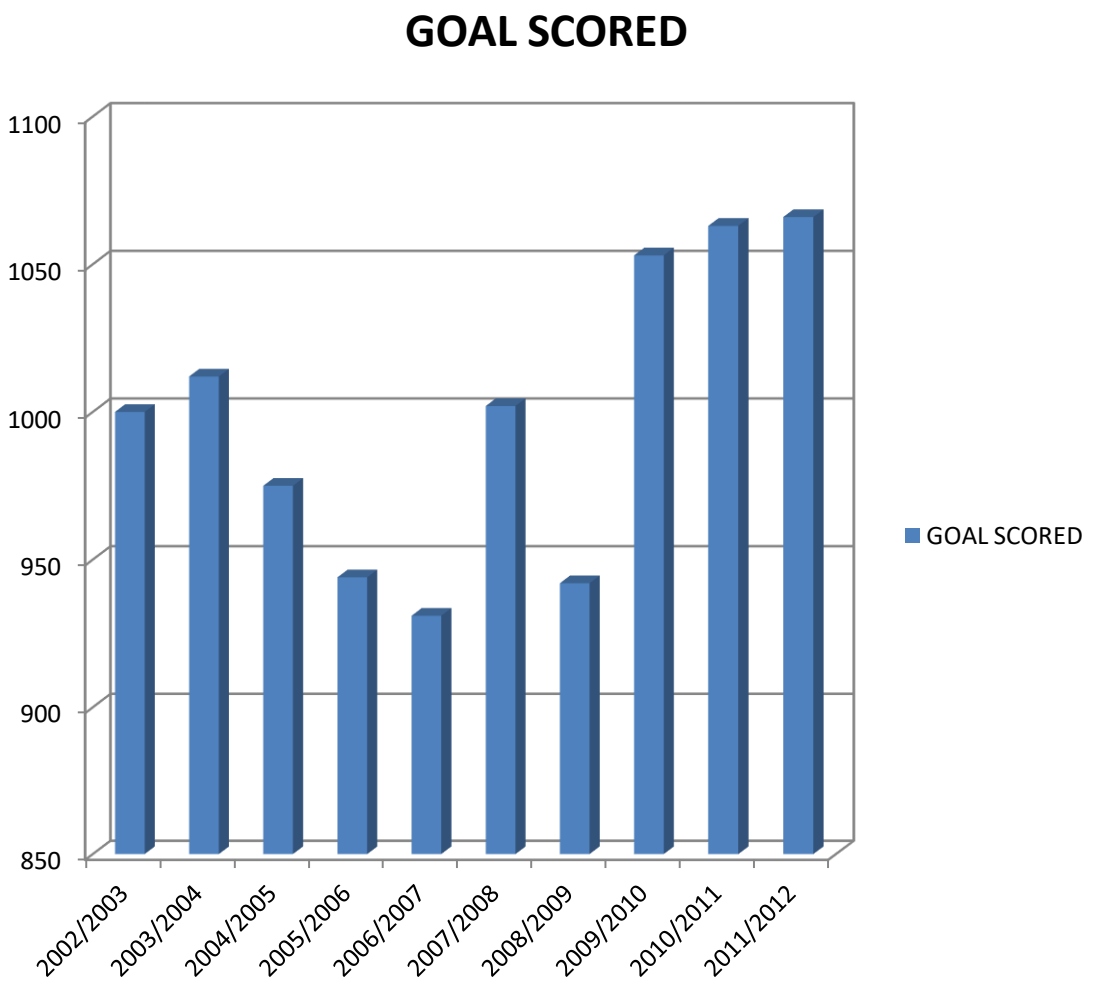

**INTERPRETATION:** The chart above show that the highest goal scored occurred in 2008/2009 and the lowest goal occurred in 2005/2006.

## LINE CHART SHOWING GOAL SCORED AGAINST TIME IN LALIGAL



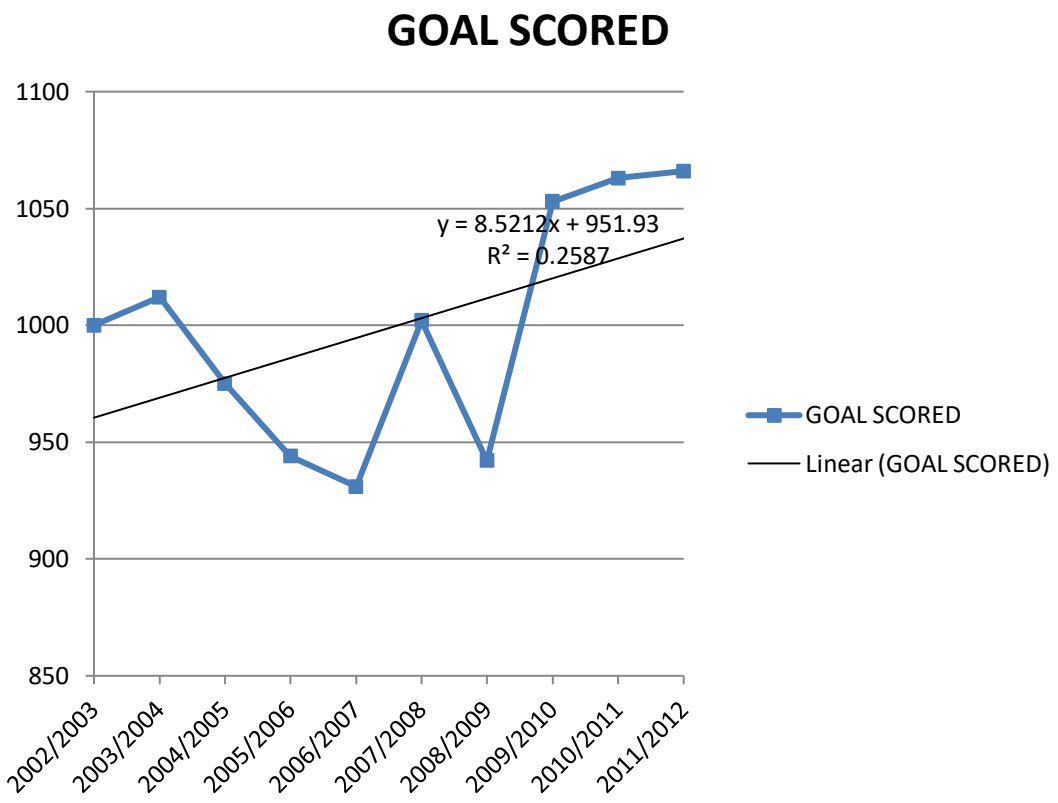
**INTERPRETATION**: The chart above shows that the highest goal scored occurred in case number 7 which is 2008/2009 and the lowest goal occurred in 2005/2006.

**SIMPLE BAR CHART SHOWING GOAL SCORED AGAINST TIME IN ENGLISH PREMIERSHIP**



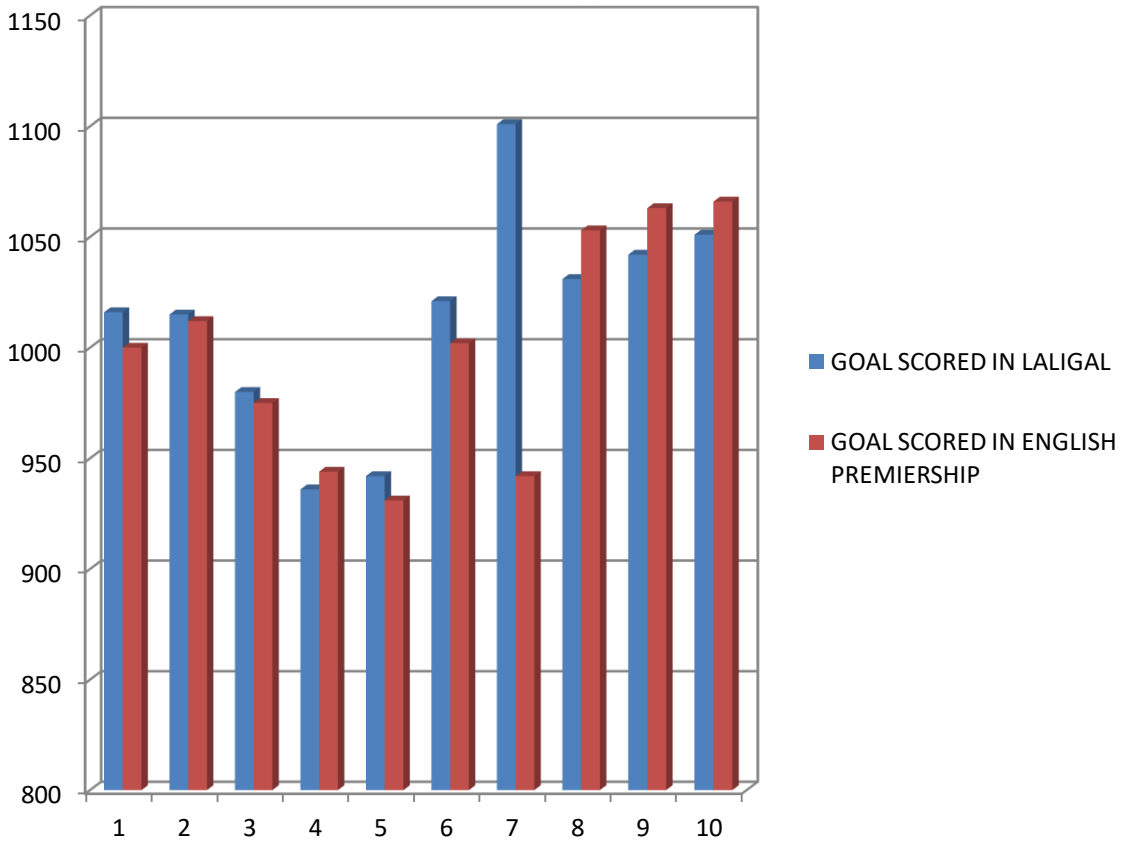
**Interpretation**: The chart above show that the highest goal scored occurred in 2011/2012 and the lowest goal occurred in 2006/2007.

**LINE CHART SHOWING GOAL SCORED AGAINST TIME IN ENGLISH PREMIERSHIP**




**INTERPRETATION:** The chart above shows that the highest goal scored occurred in case number 7 which is 2011/2012 and the lowest goal occurred in 2006/2007.

## MULTIPLE BAR CHART SHOWING GOAL SCORED IN LALIGAL AND ENGLISH PREMIERSHIP




**INTERPRETATION:** The multiple bar chart above shows that the highest goal scored in Spanish Laligal occurred in 2008/2009 season with 1101 goals and in English premiership, it occurred in 2011/2012 season with 1066 goals while the lowest goals occurred in 2005/2006

season with 936 goals in Spanish laligal and 2006/2007 season with 931 goals in English premiership league.

# CHAPTER FIVE

## 5.0 SUMMARY OF FINDINGS, RECCOMENDATION AND CONCLUSION

## 5.1 SUMMARY OF FINDINGS

Reasonable outcome cannot be drawn out from the set of data without a careful analysis of the data collected from Spanish Laligal and English Premiership league. In the previous chapter, analysis of data was carried out using various statistical tools such as Regression, Correlation, Linearity test using ANOVA, Test of hypothesis using Student T distribution and also representation of data on charts and diagrams such as simple bar chart, multiple bar chart and Line chart.

The charts and diagram show the linearity that exists between the variables, also the linear regression shows that there is no linear relationship between time and goal scored in the two leagues, likewise, correlation analysis also shows the degree of association that exist between the variables.

The correlation coefficient carried out in chapter 4 is given as 0.4868, which shows that there is a weak positive correlation between time and goal scored in Spanish

Laligal also the correlation coefficient which is 0.5087 also implies there is a strong positive correlation between time and goal scored in English premiership. The value of R2 = 0.2370 shows that about 24% of goal scored in Spanish Laligal can be explained by time leaving about 76% to be explained by other factors. Likewise the value of R2 = 0.2588 shows that about 26% of goal scored in English premiership can be explained by time leaving about 74% to be explained by other factors.

The regression model obtained in chapter 4 is given as **Y = 969.268 + 8.024x**, which implies that since the value of 'b' is positive, for every progression in 'x' which is time, there will be a corresponding increase in the number of goal scored in Spanish Laligal, it also reveal that there is no linearity between goal scored and time in Spanish Laligal. Also, **Y = 951.93 + 8.5212x**, which shows that since the value of 'b' is positive, for every progression in 'x' which is time, there will be a corresponding increase in the number of goal scored in English premiership league, it also reveal that there is no linearity between goal scored and time in English premiership league. Therefore these shows that we cannot forecast with the linear models obtained.

Student "t" analysis for comparative difference of mean between goal scored in the two leagues shows that Tcal=0.65 and Ttab=2.101, which mean Tcal <

Ttab, thereby accept my Ho and conclude that there is no significance difference between the mean goal scored in the two leagues.

## 5.2 RECOMMENDATION

Based on the result of the study, the following suggestions are recommended towards a better improvement in the sporting activity.

1. Professional officials should be appointed to officiate every big match.
2. Limitation of match played.
3. Goal line technology should be introduced by FIFA, by doing this it will reduce goal scored.
4. There should be punishment for any mis-officiating by the referee.
5. There should be total eradication of biased act in controlling the affairs of football sporting activities by FIFA.

## 5.3 CONCLUSION

If all the recommendation stated above are properly, carefully and correctly followed, I strongly believe that there will be a sustainable development

and improvement in sporting activities in the world especially football. Thank you.