2020

Effectiveness of Spatial-temporal Data Using GIS in America's Professional Sports Leagues (MLS, MLB, and NFL)

Abdullah Aleissa
Claremont Graduate University

Effectiveness of Spatial-Temporal Data using GIS in America's Professional Sports Leagues

(MLS, MLB, and NFL)

By

Abdullah Aleissa

Claremont Graduate University

2020

APPROVAL OF THE REVIEW COMMITTEE

This dissertation has been duly read, reviewed, and critiqued by the Committee listed below, which hereby approves the manuscript of Abdullah Aleissa as fulfilling the scope and quality requirements for meriting the degree of Doctor of Philosophy.

Brian Hilton, Chair

Claremont Graduate University

Clinical Full Professor in CISAT

Lorne Olfman, Member

Claremont Graduate University

Professor and Director in CISAT

June Hilton, Member

Claremont Graduate University

Adjunct Professor in Education and CISAT

Abstract

Effectiveness of Spatial-Temporal Data using GIS in America's Professional Sports Leagues (MLS, MLB, and NFL)

By

Abdullah Aleissa

Claremont Graduate University: 2020

This dissertation explores how information systems can improve the understanding of the home field advantage (HFA) in professional sports leagues in the United States. The literature related to the HFA conceptual framework and the game location factor—which represents four major impacts on teams (crowd, learning, travel, and rules)—led to an investigation into whether a relationship exists between game results and spatial, temporal, and stadium attributes. These stadium attributes include field surface type, roof type (i.e., open, fixed, and retractable), time zone, and field orientation (e.g., N/S, E/W, NE/SW) for U.S. professional sports such as soccer, baseball, and football. Winning percentage, winning streak, and losing streak were examined for their effect on game outcome. Collectively, all league games were examined to assess the effects of spatial, and temporal stadium attributes. A logistic regression (LR) analysis of the National Football League (NFL), Major League Soccer (MLS), and Major League Baseball (MLB) shows evidence of a significant relationship between spatial orientation, temporal, and stadium attributes and the game results. The LR model consider as an improvement over the base model, and the results vary from one sport to another. An IT artifact (dashboard) operationalized the proposed model based on these results. The dashboard provides team decision-makers with

information to help them understand their opponent's in the next home or away game. The artifact could be an integral part of the decision-making process for coaches and managers in game preparation and management.

Dedication

To my dear parents, Muqbel and Ibtesam

&

My lovely wife, Haneen, and my wonderful daughter, Jana

&

My whole Family members

Acknowledgments

I would like to express my deepest gratitude and appreciation to my committee chair Prof. Brian Hilton, for his infinite support, guidance, and enthusiastic encouragement throughout my PhD coursework and the dissertation process at CGU. Thank you for making it all seem possible.

Beside my advisor, I would like to thank committee members (Prof. June Hilton and Prof. Lorne Olfman) for serving on my dissertation committee and for their great support and invaluable advice. I am grateful to Prof. June Hilton for her insightful comments and for sharing with me her tremendous experience in statistics. I also show gratitude for Prof. Lorne Olfman for all the encouragement and genuine interest.

I want to acknowledge a special person, my brother Yazeed, who is very proud of my work and does not mind telling people so. I will forever be thankful and grateful for his willingness, support, and advice during this dissertation.

I am very grateful to all my friends. Specially, Yasser Khouja for his amazing sense of respect when helping me along the way of this journey. Also, I would like to thank William Kariuki for his worthy support and cooperation and time in terms of providing me all needed information.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AFC – American Football Conference

AL – American League

CRISP-DM – Cross-Industry Standard Process for Data Mining

DSR – Design Science Research

GIS – Geographic Information System

HFA - Home Field Advantage

LR – Logistic Regression

MLB – Major League Baseball

MLS – Major League Soccer

NBA – National Basketball League

NFC – National Football Conference

NFL – National Football League

NHL – National Hockey League

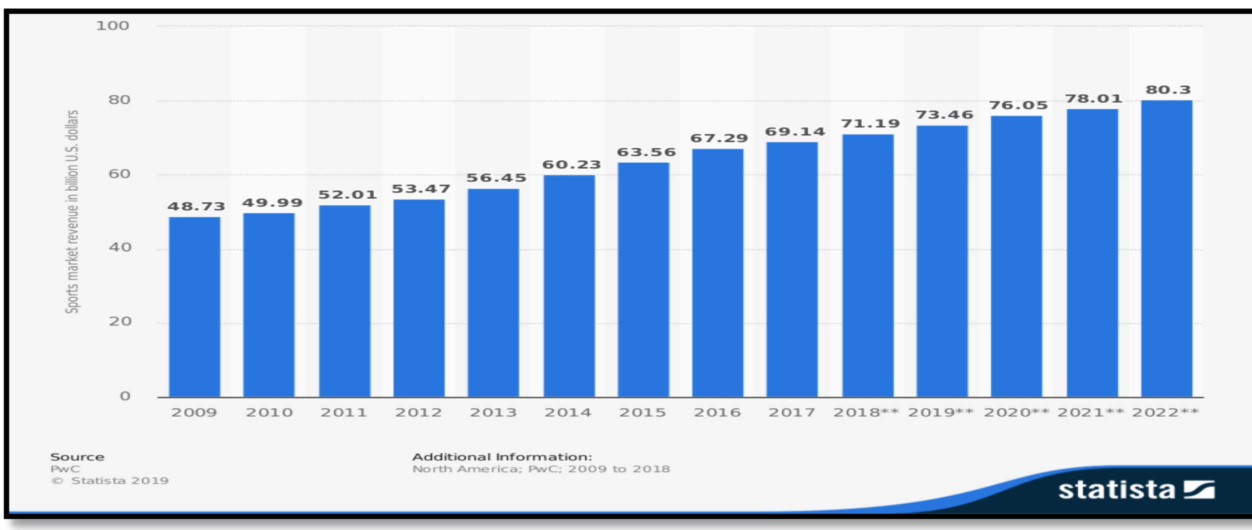
NL – National League

RTLM – Real Time Learning Machine

Chapter 1: Introduction

Sports markets are experiencing significant growth. Sports markets have developed into a worldwide industry with an extraordinary level of revenue and more players than ever (Heitner, 2016). The North American sports market is one of the largest in the world and was predicted to generate 71.8 billion dollars in revenue in 2018 (PwC.com). Figure 1.1 shows the sports market size in North America from 2009 to 2017 with predictions from 2018 to 2022. The market is a combination of revenues, media rights, sponsorships, and merchandising.

Figure 1.1
North America Sports Market Size from 2009 to 2022 (in Billions of U.S. Dollars)



Note: Retrieved from statista.com.

Sports analytics can be defined as the management of structured historical data, the application of predictive analytic models that use that data, and the use of information systems to inform decision-makers and enable their organizations to gain a competitive advantage on the field of play (Alamar, 2013).

Many industries today are acquiring more analytical approaches to decision-making; however, no other industry has the same types of analytical inventiveness as the sports industry.

Sports differ from businesses, but both domains of activity have in common the need to optimize critical resources and of course the need to win (Davenport, 2007). Different analytic methods address game and player performance, player selection, customer relationships, business management, injury prevention, and data management. Additionally, the United States hosts many major conferences for sports analytics, such as the National Sports Forum, Sports Industry Networking & Career Conference, and MIT Sloan Sports analytics conference. Information technology and analytics have had a significant impact on how sports are played, watched, and managed (Lewis, 2014).

The five major sports in the United States are football, baseball, basketball, ice hockey, and soccer (Riess, 2017). This study discusses football, baseball, and soccer. Basketball and ice hockey were eliminated from the study for two reasons, (a) both sports play on the same playing surface, and (b) all teams play inside arenas that have the same closed or fixed roof type.

MLS

MLS is a lucrative, growing sports industry in the United States; this is demonstrated in USA Today's (2017) headline: "*Major League Soccer's attendance is up, and fan interest is booming.*" MLS set a new record in 2017 with averaging 22,000 in attendance for the first time in history, ranking among the top six leagues in the world (Statista, 2020a). Forbes (2016) stated: "The average MLS team is now worth $185 million, up 18% from the 2015 valuation and up a staggering 80% from 2013 valuation and 401 percent higher than the 2008 figure of $37 million." Twenty-four teams play in the MLS, with 21 teams from the United States and three teams from Canada. The net worth of all MLS teams is more than $4,000,000,000. Furthermore, the last broadcasting deal signed by MLS for television rights was worth $90,000,000 per year and was signed by three major broadcasters: Fox, ESPN, and Univision (Smith, 2018).

Several factors inherent in the MLS training regimen and schedule are to actively work to increase team performance. Kurt Andrews (2017)—the assistant athletic trainer for the LA Galaxy—stated that almost all professional sports teams and players are now using special gear such as GPS devices or internal load monitoring systems to help sports scientists and performance teams track the daily load sustained by the athletes. The performance staff can use this information to set norms for each individual, understand how players recover from games and training sessions, and understand which players are at a higher risk of injury. In addition, this information can provide evidence of other external influences that may affect team or player performance such as crowd, stadium orientation, playing surface type, stadium and roof type.

MLB

Baseball is a sport with no fixed playing time. A regulation game consists of nine innings where the teams alternate between batting and playing defense. The away team bats first while the home team plays defense. The home team bats last in each of the nine innings, including in extra innings if the game is tied at the end of the ninth inning. Each team is allowed three outs per inning for a potential total of 27 outs per regulation game. No ties games exist in baseball (Stefani, 2008).

The MLB is a professional sports league made up of 30 teams that are evenly distributed between the American League (AL) and the National League (NL) (Statista, 2020b). Each league comprises of an east, west, and central division and each division is made up of five teams. In 2017, the league generated a total of 9.46 billion dollars, averaging approximately $315,000,000 dollars per team.

NFL

The NFL is a professional American football league that consists of 32 teams split evenly between the National Football Conference (NFC) and the American Football Conference (AFC). The NFL is one of the four main professional sports leagues in North America and is the highest professional level of American football in the world (Jozsa, 2017). The 32 NFL teams play a 16-game unbalanced schedule (Stefani, 2008). A regulation game consists of two halves and each half consist of two quarters and each quarter is 15 minutes in duration. Typically, each team has three sets of players (offense, defense, and special teams) who are substituted throughout a game.

HFA

Courneya and Carron (1992) defined the HFA as ‘‘the term used to describe the consistent finding that home teams in sports competitions win over 50% of the games played under a balanced home and away schedule’’ (p. 13). In that same article, Courneya and Carron proposed a conceptual framework of HFA that consisted of five major components: game location, game location factors, critical psychological states, critical behavioral states, and performance outcomes. Several years later, Carron and Hausenblas (1998) provided generalizations about the extent of HFA, stating that HFA is present in both professional and amateur sports. HFA is valid for individual and team sports and is generalizable across gender.

Problem Statement

During a meeting with the assistant coach of the Philadelphia Union soccer team, he proposed that any team needed to only win four out of 17 away/road games to advance to the playoffs. He raised a very interesting question “How can we win more on the road?”. His statement led this researcher to investigate team performance based on HFA and the use of information technology such as the geographic information systems (GIS) to understand this phenomenon.

This dissertation is in a three-paper format. The first paper—Chapter 2—is a review of the literature related to HFA, the game location factor, and the performance outcome factor. The second paper, Chapter 3, is a research study articulating and testing how spatial-temporal orientation and stadium location attributes such as field surface type, roof type, time zone, or field orientation affect team performance and impact HFA. Chapter 3 also investigates how winning percentage and winning and losing streak affect the HFA. Chapter 4—the third paper—is a design of an IT artifact for analyzing and responding to the findings related to the previous research questions. Lastly, Chapter 5 is a conclusion of the research and includes limitations and future research.

Research Objectives

Chapter 2 examines existing literature on HFA in sports and geospatial location and resulted in three research questions:

(1) What impact do spatial, temporal, and stadium attributes such as field surface type, roof type, time zone, or field orientation have on team performance, and subsequently, HFA?

(2) What impact does win percentage and winning / losing streak have on HFA?

(3) Can an IT artifact be developed that can be effective for analyzing and responding to the findings related to the previous research questions?

The study’s design methodology is outlined using the Cross-Industry Standard Process for Data Mining (CRISP-DM). Lastly, the study provides answers to the research questions and discusses the significance of the study results.

References

Alamar, B. C. (2013). *Sports analytics: A guide for coaches, managers, and other decision makers.* Columbia University Press.

Andrews, K. (2017). *The training session: Improving strength and conditioning at the club level.* https://www.hudl.com/blog/the-training-session-improving-strength-and-conditioning-at-the-club-level

Carron, A. V., Hausenblas, H. A., & Eys, M. A. (2005). *Group dynamics in sport.* Fitness Information Technology.

Courneya, K. S., & Carron, A. V. (1992). The home advantage in sport competitions: A literature review. *Journal of Sport and Exercise Psychology, 14*(1), 13–27. https://doi./10.4135/9781483332222.n138

Davenport, T. H., & Harris, J. G. (2007). *Competing on analytics: The new science of winning.* Harvard Business Press.

Forbes. (2016, September 07). *Forbes releases 2016 MLS team valuations* [Press release]. https://www.forbes.com/sites/forbespr/2016/09/07/forbes-releases-2016-mls-team-valuations/#308069cf4f85

Heitner, D. (2016, March 28). *Sports industry to reach $73.5 billion by 2019.* Forbes. https://www.forbes.com/sites/darrenheitner/2015/10/19/sports-industry-to-reach-73-5-billion-by-2019/#16fe62541b4b

Jozsa, F. P. (2017). *Sports capitalism: The foreign business of American professional leagues.* Routledge.

Lewis, M. (2004). *Moneyball: The art of winning an unfair game.* WW Norton & Company.

Riess, S. A. (2017). Professional Team Sports in the United States. In *Oxford Research Encyclopedia of American History*. https://doi.org/10.1093/acrefore/9780199329175.013.198

Smith, C. (2018, December 21). *Major league soccer's most valuable teams 2017.* Forbes. https://www.forbes.com/sites/chrissmith/2017/08/16/major-league-soccers-most-valuable-teams-2/#88702cb815de

Statista. (2020a). *Major soccer leagues worldwide ranked by average per game attendance in 2018/19* [Data set]. https://www.statista.com/statistics/270301/best-attended-football-stadiums-in-the-world-by-average-attendance-2010/

Statista. (2020b). *North America sports market size from 2009 to 2022 (in billion U.S. dollars)* [Data set]. https://www.statista.com/statistics/214960/revenue-of-the-north-american-sports-market/

Stefani, R. (2008). Measurement and interpretation of home advantage. In J. Albert & R. H. Koning (Eds.), *Statistical thinking in sports* (pp. 203–216). Taylor & Francis Group. https://books.google.com/books?id=ChpdantgTqQC&pg=PA203&dq=stefani+Measurement+and+interpretation+of+home+advantage.&hl=en&newbks=1&newbks_redir=0&sa=X&ved=2ahUKEwjuv6mOqI7oAhWTgnIEHRnkANAQ6AEwAHoECAAQAg

USA Today. (2017, November 17). *Major league soccer attendance is up, but TV ratings lag as U.S. soccer mulls future.* https://www.usatoday.com/story/sports/mls/2017/11/17/mls-attendance-up-tv-ratings-lag-as-us-mulls-future/107770250/

Chapter 2: Literature Review

Abstract

A sufficient amount of literature addresses the areas of sports analytics. This paper reviews the literature related to the HFA framework in general, the game location factor, and the performance outcome factor in particular. The literature review resulted in three research questions that provide a comprehensive overview of this framework and to explore whether any other factors may lead to HFA.

Overview

This paper reviews literature that discusses sports analytics and HFA. The literature search was conducted using two online databases: Google Scholar and Scopus to find the relative literature in the last fifty years. The search terms used were sports analytics, HFA, sports travel, sports crowd, sports road effect, sports familiarity, home field, or sports time zone. The literature search resulted in over 90 articles, including journal articles and conference papers. Thirty-six articles were selected for inclusion in the literature review. The selection criteria were based on the relevance of the article to sports analytics or HFA factors for professional sports. Table 2.1 shows the bibliography of the final articles.

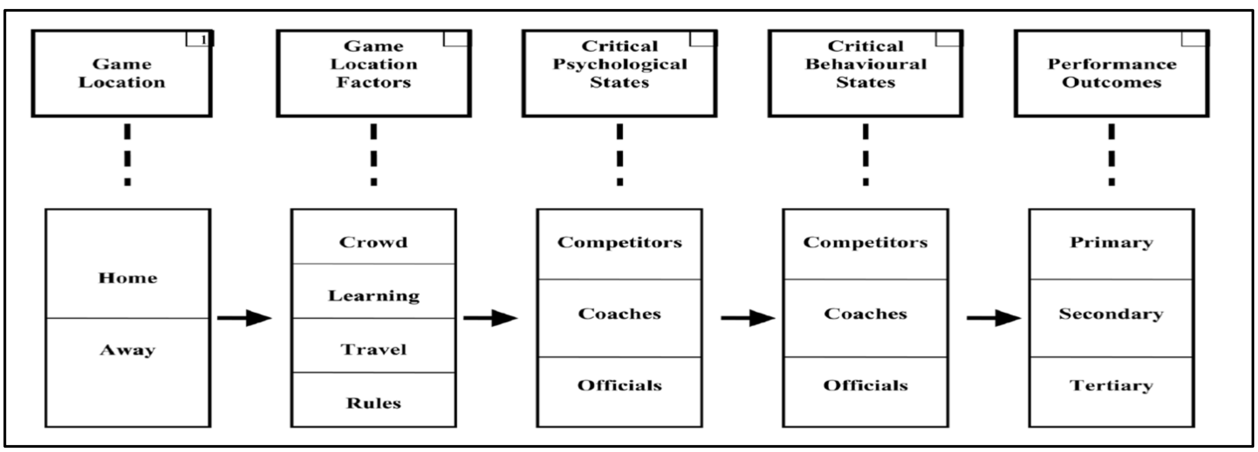
Table 2.1.
Bibliography of Final Articles

Author Last Name	Year of Publication	Source	Foucs	HA framework Relevance
Agnew	1994	International Journal of Sport Psychology	Aim to examened if the home advantage exist in major junior-A ice hocky in Canada	The crowd Effect
Bray	2000	Journal of Sport Behavior	Test intercollegiate basketball players perceptions of game location factors they believed influenced their team's performance	Psychological states Effect
Bray	2008	Journal of Applied Sport Psychology	Examine the home versus away records of individual teams in order to more fully describe team performance outcomes in relation to game location	Testing
Carron	2005	Journal of Sports Sciences	Updated framework for Home Advantage	Updated Framework
Clarke	1995	Journal of the Royal Statistical Society	Examined the home advantage for all teams in English Football League for a decade (1981-82 to 1990-91)	Testing
Courneya	1991	Human Kinetics Journal	Investigated the effects of season game number, series game number, length of home stand, length of visitor's road trip, home travel, and visitor travel on the home advantage in minor league Double A baseball	Travel and Road trip Effect
Courneya	1992	Journal of Sport and Exercise Psychology	Provide a conceptual framework for home advantage incorporate five major components	HA Framework
Courneya	1990	Journal of Sport and Exercise Psychology	Examined the rule factor in home field advantagae in softball.	Rule Factor Effect
Downward	2007	Journal of Sports Sciences	Testing the relationship between crowd size and yellow cards, because significantly higher number of yellow cards were awarded against the away team	The crowd Effect, Referee
Edwards	1979	Social and psychological viewpoints	Comparing Proffisional sports to Non-Proffisional (College) sports	Testing HFA
Gayton	1992	Perceptual and Motor Skills Journal	Examined whether the home advantage exists in individual sports	Individual sport
Goumas	2013	Journal of Sport Behavior	Investigate the role crowd size plays in home advantage and how its effect may vary worldwide	The crowd Effect
Greer	1983	Social psychology quarterly	The Impact of Social influence in Basketball Arena on HFA	The crowd Effect
Irving	1990	Journal of Sport Behavior	The effect og HFA on Performance of Baseball Pitchers	Psychological states Effect
Jamieson	2010	Journal of Applied Social Psychology	Quantify the probability of a home victory, thus only studies that included win–loss data were included in the meta-analysis	Time and Season Length
Jehue	1993	Medicine & Science in Sports & Exercise Journal	Determine the effect of time zone and game time changes on National Football League (NFL) team performance	Travel and Time zone Effect
Loughead	2003	International Journal of Sport and Exercise Psychology	Invastigating the effect of Facility familiarity and the home advantage in professional sports	Famililrity/Learning Factor
Madrigal	1999	Journal of Sport Behavior	Providing two hierarichical models explain signifcant variance in home winning percentage	Team Quality
Morley	2005	Journal of Sports Sciences	Examined the factors affecting the outcome of cricket matches played in the English one-day county cricket league	Testing
Morley	2005	Journal of sports sciences	Investigation of home advantage and other factors affecting outcomes in English one-day cricket matches	Additional Factor (One Day Tournament)
Nevill	1999	The American Journal of Sports Medicine	Examined the factors of Home Advantage and ranked the factors from the most important to less important	Testing
Nevill	2002	Psychology of Sport and Exercise	Testing the influence of crowd noise upon refereeing decisions in football	The crowd Effect, Referee
Pace	1992	Canadian Journal of Sport Sciences	Examine the relative contributions of various travel related variables to visiting team success in the National Hockey League	Travel Effect
Pollard	2002	Journal of Sports Sciences	The effect of moving to new stadium on the home field advantage because the familiarity with the local playing facility in three different sports	Famililrity, New Home
Pollard	2017	International Journal of Performance Analysis in Sport	Use a multivariate approach to investigate variations in home advantage between different team sports and between different countries both for men's and women's competition	Testing
Pollard	2002	Journal of Sports Sciences	Provide Evidence of a reduced home advantage when a team moves to a new stadium	Famililrity/Learning Factor
Poulter	2009	Journal of Sports Sciences	Investigate home advantage at a team and individual player level as well as determine the effect of player nationality on home advantage	Additional Factor (Foreign Player)
Salminen	1993	Perceptual and Motor Skills Journal	See how much the audience effect on the home advantage in three different sports	The Crowd Effect
Schwartz	1977	Social forces	This investigation confirms the existence of a home advantage in organized sports	Testing HFA
Smith	2000	Human Kinetics Journal	At least for some professional sports, team travel can exert a very small influence on the outcome of the contest even after the quality of the teams competing is controlled	Travel and Road trip Effect
Snyder	1985	Sociology of sport journal	Indicate that home teams win 66% of their games in college basletball	Testing HFA
Staufenbiel	2015	Journal of Sports Sciences	Examined soccer coaches' expectations, goal setting and tactical decisions in relation to game location	Psychological states Effect
Steenland	1997	American Sleep Disorders Association and Sleep Research Society	The effect of travel and rest on teams and players performance	Travel Effect
Stefani	2008	Statistical thinking in sports	The importance of Psychological and Physiological factors in Home Field Advantage	Psychological states Effect, Physiological facor
Terry	1998	Journal of Science and Medicine in Sport	Investigate the relationship between game location and precomputation psychological states	Psychological states Effect
Unkelbach	2010	Human Kinetics Journal	Propose that crowd noise correlates with the criteria referees have to judge	The crowd Effect, Referee

HFA Framework

Courneya and Carron (1992) proposed a conceptual framework for HFA (see Figure 2.2) that incorporates five major components: game location, game location factors, critical psychological states, critical behavioral states, and performance outcome. The game location component represents the site for the game (home games versus away games). The game location factors represent four major impacts on teams: crowd, learning, travel, and rules. All four game location factors are considered to influence teams. Critical psychological states influence critical behavioral states, and both of the latter two components have the same three actors involved in the outcome: coaches, competitors, and officials. Finally, Courneya and Carron noted that three levels of performance outcomes exist: primary, secondary, and tertiary.

Figure 2.1.
HFA Conceptual Framework



The percentage of home field wins for the major team sports are as follows.

- MLB: 53.5%
- NFL: 57.3%
- National Hockey League (NHL): 61.1%
- National Basketball Association (NBA): 64.4%

- MLS: 69%.

Carron and Hausenblas (1998) used the framework and literature review to provide additional generalizations about HFA in sports. Carron and Hausenblas found that the HFA is present in both professional and amateur sports and in individual and team sports, as well as across gender. Soccer and basketball had the highest regular-season home field advantage measured by the fraction of home wins minus the fraction of home losses, followed by football, ice hockey, and baseball (Stefani, 2008).

This literature review provides a comprehensive overview of the HFA framework, investigates the advantage of playing at home, and analyzes if any other factors may lead to HFA. The literature on HFA examines the four fundamental elements of game location factors, which represents four major conditions that differently influence teams playing at their own venue versus an opponent's venue. The four elements are crowd factor, learning and familiarity factor, travel factor, and rules factor. In addition to that, the performance outcomes factor represents primary, secondary, and tertiary performance.

Crowd Factor

Crowd behavior has been investigated to discover why the crowd factor may affect the HFA. Agnew and Carron (1994) studied several factors connected with the presence of crowds at major junior-A ice hockey games in Canada. The results indicated that HFA grows as crowd density increases. Similarly, Goumas (2013) investigated the role crowd size plays in HFA and how the effect of crowd size may vary worldwide. The study results indicated that home field advantage increases by 1.5% per a 10% increase in crowd size. Nevill et al. (2002) found that English and Scottish soccer leagues had higher home-winning percentages when the crowd size was large.

Home team performance is better than the opposing team during both normal and booing crowd behavior conditions (Greer, 1983). Salminen (1993) studied three different sports in Finland: soccer, ice hockey, and basketball. Salminen's study results indicated that the home field advantage of 59% was only a little higher than the average 51%. In contrast with previous findings, Morley and Thomas (2005) found that the crowd effect is mostly insignificant in English one-day cricket matches; however, the model used did not provide any conclusive evidence about isolating and explaining any home field advantage effect. This study has a limitation on data quality for two reasons: (1) the data are incomplete and unbalanced (2) the informational content of reverse fixture comparisons when paired matches are played in separate seasons.

Travel Factor

Only a few studies have examined travel as a potential factor to home field advantage before Courneya and Carron introduced the home field advantage framework in 1992. Snyder and Purdy (1985) studied college basketball teams and found a home-winning percentage of 58.8% when visiting teams traveled less than 200 miles to their opponent’s venue. The home field advantage increased to 84.6% when visiting teams traveled more than 200 miles for a game. Similarly, Pollard (1986) examined 3,500 professional English soccer matches and found that the home field advantage for the home team increased to 64.3% when the visiting team traveled for 200 miles or more. Steenland and Deddens (1997) studied 8,495 regular season games in the NBA over eight seasons (1987–1988 through 1994–1995) to see the effect of travel and rest on team and player performance. The study found that performance for both home and visiting teams improved with more than 1 day between games. Team performance was negatively affected with only one day between games, perhaps due to lack of sleep or lack of time for musculoskeletal recovery. The visiting team—who is on the road and subject to repeated travel—is affected more than the home team.

Learning and Familiarity Factor

The learning and familiarity factor indicate that opponents are generally more familiar with their own venue. The role of learning and familiarity in contributing to HFA received limited attention until 1992, when the home field advantage framework was introduced. Pollard (2002) investigated the effect of moving to a new stadium on the HFA in baseball, basketball, and ice hockey. Pollard asserted that the HFA during the first season in a new stadium was significantly less than the HFA in the final season in the old stadium; the reduction was evident in all three sports. Similarly, Loughead et al. (2003) studied 57 teams that relocated to a new venue from the NBA (1991–2000), NHL (1982–2000), and the English and Scottish Professional Football Associations (1988–2000). Loughead et al. found that teams had an overall home-winning percentage of 55.2% before relocating to a new venue. Immediately after relocating, the home-winning percentage decreased to 53.9%. However, it has been suggested that facility familiarity is not a contributor to the home advantage (Moore & Brylinsky, 1993). Loughead et al.'s study had limitations in several areas; for example, the sample size was only 18 basketball games and the study was hampered by a lack of experimental control.

Rules Factor

The rules factor refers to the concept that some sport rules may favor the home team, such as last line change in ice hockey and in baseball, where the home team has the opportunity to submit its player roster last in addition to batting last. The rule differences based on game location that may affect HFA is limited to certain sports; however, batting last does not provide a

HFA (Courneya & Carron, 1990). No differences were found for the moderating variables of ability level, gender, or time of season. No studies have examined the impact of rules on the home field advantage since 1992 (Carron et al., 2005).

Performance Outcomes

Performance outcomes influenced by game location can be measured in three different levels: primary, secondary, and tertiary (Courneya & Chelladurai, 1991).

The primary measures represent the first stage of performance outcome in competitions and most closely reflect fundamental skill execution, such as batting average in baseball, free throw percentage in basketball, and penalties per game in soccer. The secondary measures express the intermediate stage of performance outcomes in competitions. Secondary measures usually reflect the scoring necessary to win a contest—points scored, goals allowed in basketball, football, or soccer—or subtle variations such as runs batted in and earned run average in baseball. The tertiary measures are the traditional outcome measures. Tertiary measures indicate the final outcome of the contest, such as wins and losses, point differential, and the ratio of the final score. All three levels of performance measures have been investigated in home advantage research (Edwards & Archambault, 1979; Irving & Goldstein, 1990; Schwartz & Barsky, 1977; Nevill & Holder, 1999).

Geospatial Analysis

GIS is a system designed to collect, store, handle, analyze, and present all types of geographical data. The key word to this technology is geography, which means that some portion of the data is spatial. In GIS, each record and digital object are associated with a geographical location. GIS may provide considerable useful information to support decision-making (Othenin-Girard et al., 2015).

GIS is used to identify characteristics, patterns, and movements of players and teams on the field of play. Additionally, GIS is actively used in sports, from choosing stadium locations to managing security at sporting events (Narain, 2017). GIS can enhance the study of sports at any scale from global to local; for example, GIS can be used to choose a city to host the next Olympic Games or to track a soccer player's location throughout a game. GIS is actively used in football, soccer, basketball, tennis, and other sports for accurate athlete performance analysis.

Real Time Learning Machine (RTLM)

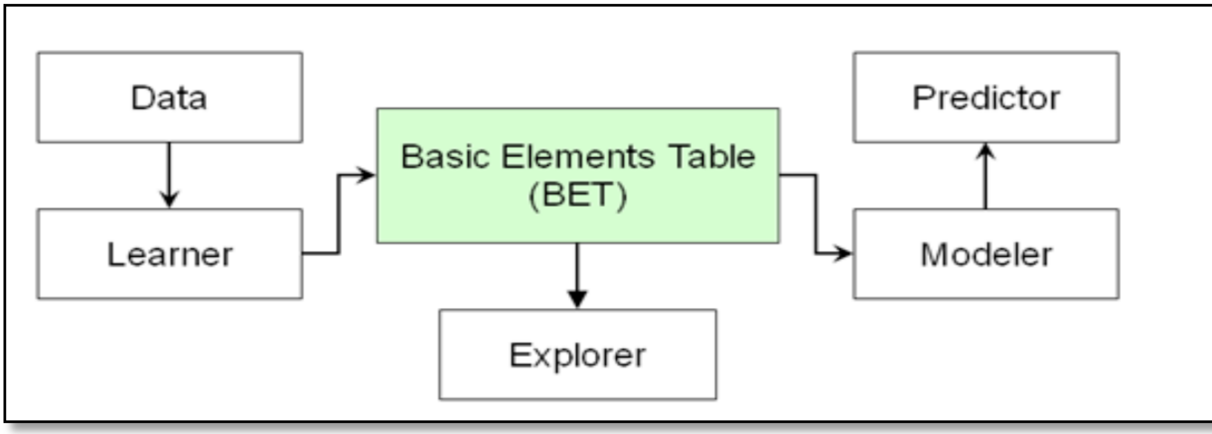
A RTLM is a method that can be easily upgraded to accept data as they are generated and to flexibly deal with changes in how data are being processed. Data mining explains the past and predicts the future by exploring and analyzing data (Sayad, 2011). The term "real-time" is used to describe how well a data mining algorithm can promptly accommodate an ever-increasing data load. The use of RTLM with traditional data mining methods enables real-time data mining.

RTLM Characteristics

The tasks assigned to each of the four real-time components are as follows:

- Learner: updates the basic elements table using the data in real time.
- Explorer: does univariate and bivariate statistical data analysis using the basic elements table in real time.
- Modeler: construct models using the basic elements table in real time.
- Predictor: uses the models for prediction in real time.

Figure 2.2.
Real Time Learning Machine



Note: Retrieved from Sayad, 2011.

Conclusion

A substantial amount of research has been accomplished in the areas of sports analytics; however, no prior studies address HFA in sports in relation to spatial-temporal and stadium location attributes such as field surface type, roof type, time zone, or field orientation for soccer, baseball, and football in the United States. Additionally, no prior studies address the effect of winning percentage and winning and losing streak on HFA.

This literature review is limited as other publications may have emerged after the initial articles were gathered. The following research questions were identified as a result of this literature review and will be addressed in the second paper.

RQ1. What impact do spatial, temporal, and stadium attributes such as field surface type, roof type, time zone, or field orientation have on team performance, and subsequently, HFA?

RQ2. What impact does win percentage and winning / losing streak have on HFA?

RQ3: Can an IT artifact be developed that can be effective for analyzing and responding to the findings related to the previous research question?

References

Agnew, G. A., & Carron, A. V. (1994). Crowd effects and the home advantage. *International Journal of Sport Psychology. 25*(1), 53–62.

Carron, A. V., Loughhead, T. M., & Bray, S. R. (2005). The home advantage in sport competitions: Courneya and Carron's (1992) conceptual framework a decade later. *Journal of Sports Sciences, 23*(4), 395–407. https://doi.org/10.1080/02640410400021542

Courneya, K. S., & Carron, A. V. (1990). Batting first versus last: Implications for the home advantage. *Journal of Sport and Exercise Psychology, 12*(3), 312–316.

Courneya, K. S., & Carron, A. V. (1992). The home advantage in sport competitions: A literature review. *Journal of Sport and Exercise Psychology, 14*(1), 13–27. https://doi.org/10.1123/jsep.12.3.312

Courneya, K. S., & Chelladurai, P. (1991). A model of performance measures in baseball. *Journal of Sport & Exercise Psychology, 13*(1).

Edwards, J., & Archambault, D. (1979). The home field advantage. In J. H Goldstein (Ed.), *Sports, games, and play: Social and psychological viewpoints* (pp. 409–438).

Goumas, C. (2013). Modelling home advantage in sport: A new approach. *International Journal of Performance Analysis in Sport, 13*(2), 428–439. https://doi.org/10.1080/24748668.2013.11868659

Greer, D. L. (1983). Spectator booing and the home advantage: A study of social influence in the basketball arena. *Social Psychology Quarterly, 46*(3), 252–261. https://doi.org/10.2307/3033796

Irving, P. G., & Goldstein, S. R. (1990). Effect of home-field advantage on peak performance of baseball pitchers. *Journal of Sport Behavior, 13*(1), 23.

Loughead, T. M., Carron, A. V., Bray, S. R., & Kim, A. (2003). Facility familiarity and the home advantage in professional sports. *International Journal of Sport Psychology and Exercise Psychology, 1*(3), 264–274. https://doi.org/10.1080/1612197X.2003.9671718

Moore, J. C., & Brylinsky, J. A. (1993). Spectator effect on team performance in college basketball. *Journal of Sport Behavior, 16*(2), 77–84.

Morley, B., & Thomas, D. (2005). An investigation of home advantage and other factors affecting outcomes in English one-day cricket matches. *Journal of Sports Sciences, 23*(3), 261–268. https://doi.org/10.1080/02640410410001730133

Narain, A. (2017, September 15). *Sports on course to become 'geo-tech' with GIS, GPS, and drones.* Geospatial World. https://www.geospatialworld.net/blogs/gis-gps-and-drones/

Nevill, A. M., & Holder, R. L. (1999). Home advantage in sport. *Sports Medicine, 28*(4), 221–236. https://doi.org/10.2165/00007256-199928040-00001

Nevill, A. M., Balmer, N. J., & Williams, A. M. (2002). The influence of crowd noise and experience upon refereeing decisions in football. *Psychology of Sport and Exercise, 3*(4), 261–272. https://doi.org/10.1016/S1469-0292(01)00033-4

Othenin-Girard, C., Guillemette, M. G., Foley, E., & Caron, C. (2015). When competitive intelligence meets geospatial intelligence. *International Journal of Business Intelligence Research (IJBIR), 6*(2), 33-54. https://doi.org/10.4018/IJBIR.2015070103

Pollard, R. (1986). Home advantage in soccer: A retrospective analysis. *Journal of Sports Sciences, 4*(3), 237–248. https://doi.org/10.1080/02640418608732122

Pollard, R. (2002). Evidence of a reduced home advantage when a team moves to a new stadium. *Journal of Sports Sciences, 20*(12), 969–973. https://doi.org/10.1080/026404102321011724

Salminen, S. (1993). The effect of the audience on the home advantage. *Perceptual and Motor Skills, 76*(3), 1123–1128. https://doi.org/10.2466/pms.1993.76.3c.1123

Sayad, S. (2011). *Real time data mining.* Self-Help Publishers.

Schwartz, B., & Barsky, S. F. (1977). The home advantage. *Social Forces, 55*(3), 641–661. https://doi.org/10.1093/sf/55.3.641

Snyder, E. E., & Purdy, D. A. (1985). The home advantage in collegiate basketball. *Sociology of Sport Journal, 2*(4), 352–356. https://doi.org/10.1123/ssj.2.4.352

Steenland, K., & Deddens, J. A. (1997). Effect of travel and rest on performance of professional basketball players. *Sleep, 20*(5), 366–369. https://doi.org/10.1093/sleep/20.5.366

Stefani, R. (2008). *Measurement and interpretation of home advantage.* In J. Albert & R. H. Koning (Eds.), *Statistical thinking in sports* (pp. 203–216). Taylor & Francis Group. https://books.google.com/books?id=ChpdantgTqQC&pg=PA203&dq=stefani+Measurement+and+interpretation+of+home+advantage.&hl=en&newbks=1&newbks_redir=0&sa=X&ved=2ahUKEwjuv6mOqI7oAhWTgnIEHRnkANAQ6AEwAHoECAAQAg

Terry, P. C., Walrond, N., & Carron, A. V. (1998). The influence of game location on athletes' psychological states. *Journal of Science and Medicine in Sport, 1*(1), 29–37. https://doi.org/10.1016/S1440-2440(98)80006-6

Chapter 3: Research Study and Analysis

The Impact of Spatial, Temporal, and Stadium Attributes on Home Field Advantage: An Exploratory Data Analysis

Abstract

This paper offers a statistical analysis to investigate HFA in the NFL, MLB, and MLS by examining spatial, temporal, and stadium attributes such as field surface type, roof type, time zone, and field orientation. In addition, this paper explores if winning percentage and winning / losing streak has any impact on HFA. This study used the CRISP-DM to guide the research process. The LR model was statistically significant and the results vary from one sport to another.

Introduction

This paper offers a statistical analysis to investigate HFA in the MLS, MLB, and NFL. One of the significant issues in this analysis was finding the data in a form that can be gathered, stored, and analyzed. This paper aims to investigate the implications of spatial, temporal, and stadium attributes such as field surface type, roof type, time zone, field orientation, winning percentage, and winning / losing streak for the MLS, MLB, and NFL.

Courneya and Carron's (1992) proposed HFA framework incorporates five major components: game location, game location factors, critical psychological states, critical behavioral states, and performance outcome. However, this study examines the game location factors, which represent four major indictors that may influence teams playing at their own venue versus an opponent's venue. Game location factors include the crowd factor, learning and familiarity factor, travel factor, and rules factor. This study also investigates the performance outcomes factors, which represent primary, secondary, and tertiary levels of performance. Tertiary measures are the traditional outcome measures that indicate the final outcome of the contest.

Research Methodology

The proposed extension of the HFA framework (see Figure 3.1) illustrates the factors that may help the visiting team win or avoid losing against the home team. The proposed extension of the HFA framework groups variables that share similar features. The variables of stadium playing surface type, stadium roof type, and stadium orientation are organized as stadium location attributes, while the time zone variable is categorized as the spatial-temporal attribute. Lastly, winning percentage, winning streak, and losing streak are grouped as results attributes.

The CRISP-DM methodology guided the statistical analysis process. CRISP-DM was introduced by Wirth and Hipp (2000); it has been used in multiple industries. The CRISP-DM is perhaps the best-known approach for data mining (Harper & Pickett, 2006); 51% of data mining practitioners describe CRISP-DM as their main data mining methodology. The CRISP-DM breaks the process of data mining into six major phases: business understanding, data understanding, data preparation, modeling, evaluation, and deployment (see Figure 3.2).

Figure 3.1.
Proposed extension of HFA Framework

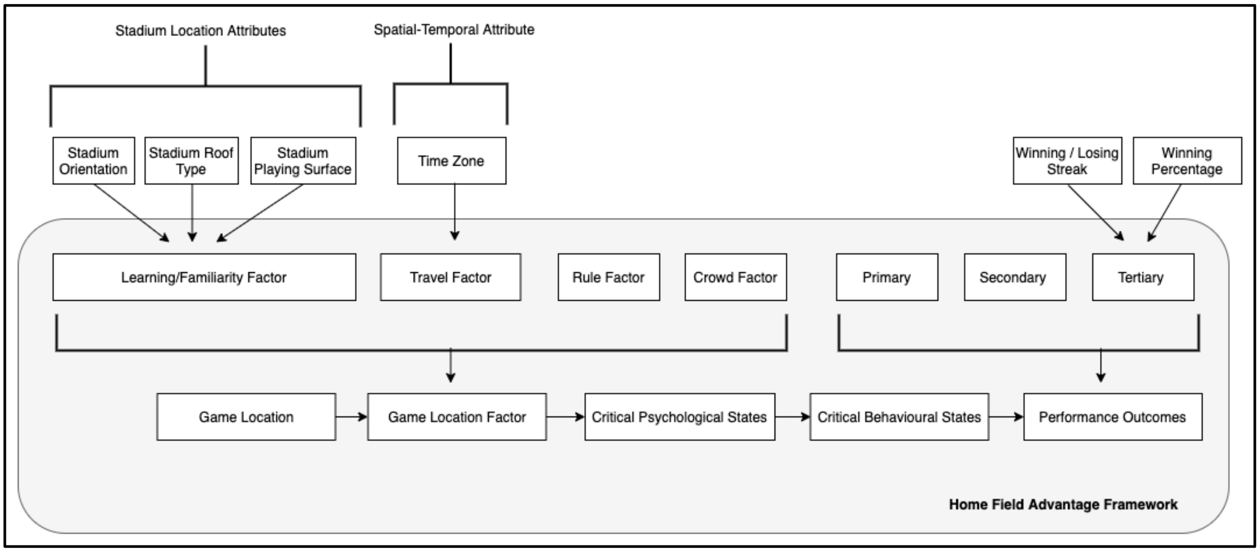
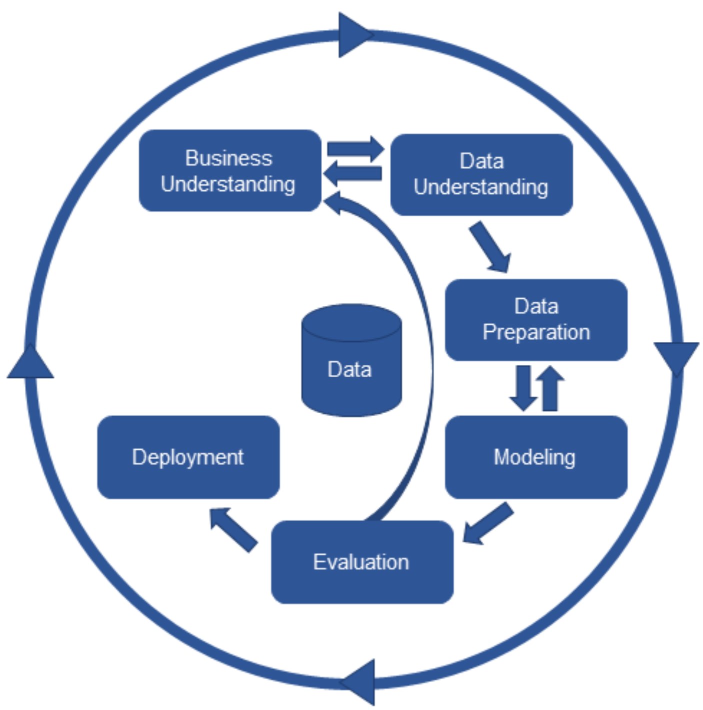


Figure 3.2.
CRISP-DM Process Diagram



Note: Retrieved from Wirth and Hipp (2000).

Major CRISP-DM Stages

Business understanding. Business understanding explains what a group aims to accomplish from a business perspective. The goal of this stage is to uncover important factors that could influence the outcome of the project. Business understanding can be divided into four subcategories.

- Determination of business objective: The business objective is to help teams win or avoid losing during away games in the NFL, MLB, and MLS.
- Assess the current situation: The NFL, MLB, and MLS have a HFA averaging 60% of any home field match.
- Determine data-mining goals: Data-mining goals should address the impact of stadium spatial-temporal attributes, winning percentage, and winning or losing streak on team performance, game results and on the HFA framework components in general.
- Produce project plan: The initial selection of tools and techniques are as follows:
    - Inputs (data source): NFL, MLB, MLS, Esri ArcMap, Opta, Other
    - Tools: Microsoft Excel, R-Studio, SPSS, Power BI, Python
    - Outputs: web application, charts, CSV files

Data Understanding. This stage requires acquisition of the data listed in the project resource. Data understanding is broken into four sections.

- Initial data collection: Data is collected from different sources (see Table 3.1). The data collected included game information, team information, travel direction, distance, and all stadium information.

- Describe the data: Summarize and describe the data that has been acquired in the study.
- Explore the data: Data exploration can be conducted using simple statistical tests in Microsoft Excel, such as the mean, min, max, and sum. Relationships between attributes can also be tested.
- Verify data quality: The quality of the data can be examined by checking if the data has missing values and identifying if the data are correct and complete.

Data Preparation. This stage involves all action needed to create the final dataset.

- Select the data: Most of the variables and attributes in this study were available online or calculated in Esri ArcMap. The travel itinerary variable in the MLS dataset was eliminated from the study due to it not being available in the MLB and NFL datasets.
- Clean the data: A number of variables required updates or required that some information be deleted, such as the rows for postponed matches.
- Construct required data: New fields were created to perform analysis on travel direction, travel distance, away team time zone difference, winning percentage, winning streak, and losing streak.
- Integrate the data: Data obtained from different sources was updated or added to the existing dataset to match the format of the date field or added to the game result field among all three datasets.

Modeling. This stage is usually conducted in multiple iterations. It is divided into four parts:

- Select modeling technique: SPSS binary LR analysis requires the data to be fully instantiated (data types are known) before execution. The dependent variable in this study will be the game result (win or loss) binary, which is represented using either 0 or 1.
- Generate test design: The analysis is designed to cover all three sports in three different scenarios. The first scenario for a three-season analysis. The second scenario is for a single season analysis. Finally, one-team analysis is conducted by testing the best record team, worst record team, and average record team.
- Build model: The model should examine alternative modeling algorithms and parameter settings.
- Assess the model: Fine tuning of the model settings according to an initial assessment of the model performance.

The remaining stages of CRISP-DM—evaluation and deployment—will be discussed in detail in Chapter 4.

Data Source

The Philadelphia Union—a professional soccer team that competes in the MLS as a member club—provided statistical data regarding their club and all other MLS teams. Additionally, relevant data was collected relating to various variables from different sources (see Table 3.1).

Table 3.1.
Data Sources

Variable	Source
Game Results	MLS: https://matchcenter.mlssoccer.com/schedule MLB: http://mlb.mlb.com/mlb/schedule/index.jsp?src=main#date=03/29/2018 NFL: https://www.nfl.com/scores
Away Team Itinerary Information	MLS: Provided by Philadelphia Union Team MLB: N/A NFL: N/A
Away Team Travel Direction	MLS, MLB and NFL: Calculated in ArcMap 10.6
Away Team Time Zone Difference	MLS, MLB and NFL: Calculated in ArcMap 10.6
Distance	MLS, MLB and NFL: Calculated in ArcMap 10.6
Home and Away Team Conference	MLS: https://www.mlssoccer.com/standings MLB: https://www.mlb.com/team NFL: https://www.nfl.com/teams
Home and Away Team Division	MLS: N/A MLB: https://www.mlb.com/team NFL: https://www.nfl.com/teams
Home and Away Team City, State, Country	MLS: https://www.mlssoccer.com MLB: https://www.mlb.com/team NFL: https://www.nfl.com/teams
Home and Away Team Time Zone	MLS: http://harvardsportsanalysis.org/2015/05/does-travel-distance-effect-results-in-mls/ MLB: https://noahveltman.com/mlbschedule/ NFL: Obtained in ArcMap 10.6
Home and Away Team Stadium Roof Type	MLS, MLB and NFL: Collected from different sources Appendix A
Home and Away Team Stadium Playing Surface	MLS, MLB and NFL: Collected from different sources Appendix A
Home and Away Team Stadium Orientation	MLS: Obtained by using GIS procedure MLB: http://www.baseball-almanac.com/stadium/ballpark_NSEW_AL.shtml NFL: Obtained by using GIS procedure

The study also included additional data provided by the Philadelphia Union (travel itinerary information) while the remaining variables (all team information, travel direction, distance, and all stadium information) were collected. See Appendix A for all stadium information resources.

Variables

Game result. This variable represents the outcome (win, loss, and tie) for each game in the MLS, MLB, and NFL, along with winning percentage, winning streak, and losing streak.

Away team itinerary information. This variable represents departure date, departure city, flight number, departure time, arrival time, return date, return city, flight number, departure time, and arrival time. This information is only for the Philadelphia Union data set. This information was not included for the other sports, as the away team travels only one night before any game most of the time.

Away team travel direction. This variable represents the travel direction for the away team (east to west or west to east) using GIS to determine the flight direction for the visiting team.

Away team time zone difference. This variable represents the change in time zones traveled [-3 to +3] by the away team. The data contains the time zones in the United States and Canada (EST, CST, MST, and PST); the data does not include AKST and HST, as no professional teams are present in Alaska or Hawaii.

Distance. This variable represents the Euclidean distance (in miles) from the away team stadium to the home team stadium. This variable was calculated in ArcMap 10.6 by using the point distance tool for each stadium for the MLS, MLB, and NFL. This tool creates a table with distances between two sets of points. Distances from all input points (the stadium locations) to all near points (all other stadiums in the dataset) are calculated.

Home and away team information. This variable represents all team information (city, state, country, time zone, team conference, and team division) and was collected from each team's website and from the professional association for each sport.

Home and away stadium information. This variable represents all stadium information (stadium orientation, stadium roof type, and stadium playing surface) where stadium orientation was determined using GIS. The stadium was located on a satellite map and assigned an orientation (N/S, E/W, etc.). The stadium roof type and stadium playing surface were collected from each team's website.

Data Set

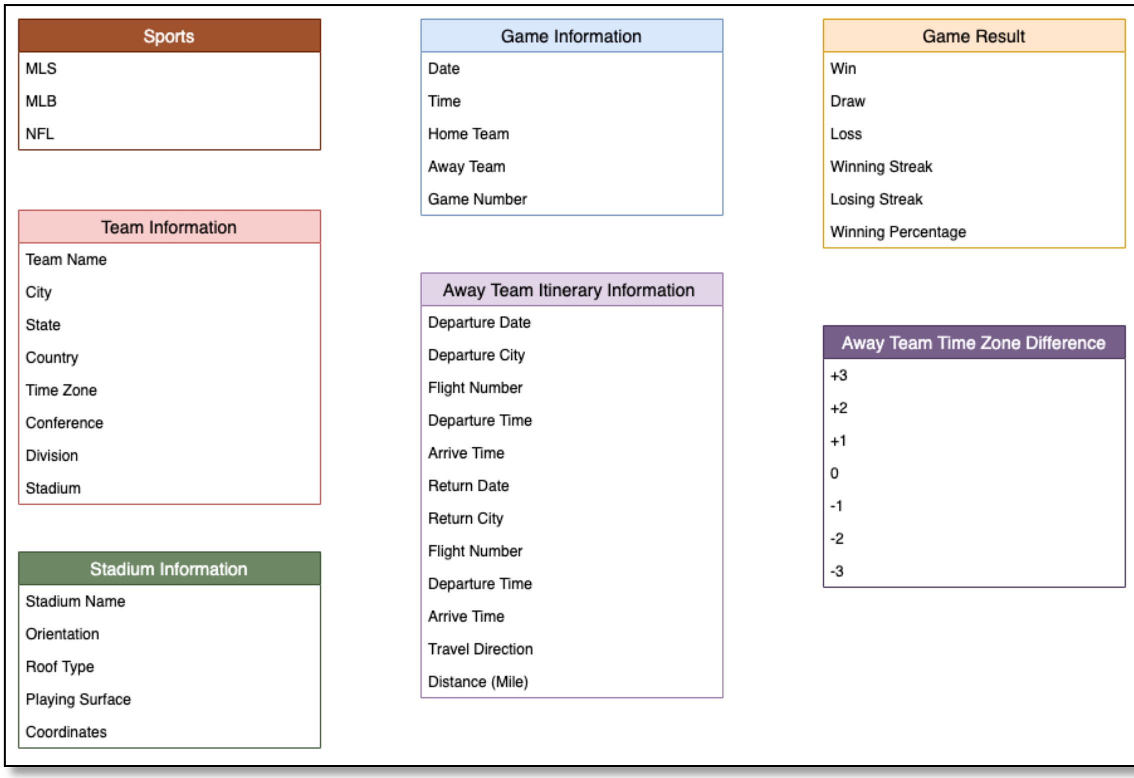
The data collection process resulted in three data sets: three seasons of MLS from 2016–2018 (N=1106), three seasons of MLB from 2016–2018 (N=7290), and three seasons of NFL from 2016–2018 (N=768).

For NFL and MLB datasets, the same number of teams competed in all three seasons: 32 teams in NFL and 30 teams in MLB. Twenty teams competed in the MLS in the 2016 season. In the 2017 season 22 teams competed after Atlanta United and Minnesota United FC joined the league. In the 2018 season, 23 teams competed after Los Angeles FC joined the league.

Data Model

A conceptual high-level diagram of the sports analytics data model is shown in Figure 3.3. This model establishes the entities, and their attributes.

Figure 3.3.
Data Model (Conceptual Design)



Preliminary Results

An initial review of the MLB, NFL American Conference (AC) and National Conference (NC), and the MLS Eastern Conference (EC) and Western Conference (WC) team performance and division performance was performed. Initial findings indicated significant correlations between the winning team, time zone, stadium orientation, and team conference. See Appendix B for all correlation results.

Analysis

This section analyzes how the variables correlate with each other in a holistic way using LR. LR was proposed as an alternative analysis method to Fisher's method (linear discriminant analysis) in the late 1960s and early 1970s (Cabrera, 1994) and became routinely available in statistical packages in the early 1980s. LR is a powerful analytical technique that can be used when the outcome variable is dichotomous (Peng et al., 2002). Like all regression analysis, LR is a predictive analysis method used to describe data and explain the relationship between one dependent binary and one or more nominal, ordinal, interval, or ratio-level independent variables.

The Dependent Variable (DV) in this study is a binary value (0 or 1) representing the game result (win or lose). A draw was eliminated for several reasons.

1. No tie games exist in baseball; the game must end with a winning team.
2. The number of tie games in the NFL (2012–2018) is very limited; less than 1% of games in eight full seasons ended as a tie game. The data set used in this study (2016–2018) included 0.52% tie games. This percentage represents only four games out of 768.
3. In soccer games, the draw results at your home game are considered as losing 2 points (out of a possible 3 for a win). This study aims to help the away team win more on the road.

Analysis Results. SPSS [1] is a statistical application used to analyze the data. The data sets were also analyzed in R [2], resulting in the same findings. The following are the analytical findings for the NFL, MLS, and MLB 2016–2018 seasons.

[1] IBM SPSS Statistic (version 25), software platform offers advanced statistical analysis, a vast library of machine learning algorithms, text analysis, open source extensibility, integration with big data and seamless deployment into applications

[2] R (version 3.5.3) is a programming language and free software environment for statistical computing and graphics supported by the R Foundation for Statistical Computing

NFL.

- DV is home team win. A LR was performed to ascertain the effects of Home team time zone is central standard time (H_CST), Away team stadium orientation is east to west (AO_EW), and Away team roof type is fixed(AR_F) on the likelihood if the Home_Team_Win. The LR model was statistically significant and all IVs were significant: $X^2(3) = 26.121$, $p < 0.001$. The model explained 3.3% (4.5%) of the variance in Home_Team_Win using Cox & Snell R^2 (Nagelkerke R^2) and correctly classified 60.8% of cases (see Figure 3.4).
- DV is away team wins. A LR was performed to ascertain the effects of H_CST, AO_EW, and AR_F on the likelihood if the Away_Team_Win. The LR model was statistically significant and all IVs were significant: $X^2(3) = 25.833$, $p < 0.001$. The model explained 3.3% (4.5%) of the variance in Away_Team_Win using the Cox & Snell R^2 (Nagelkerke R^2) and correctly classified 61.3% of cases (see Figure 3.5).

Figure 3.4.
SPSS Results NFL 2016–2018 Home Team Win

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	26.121	3	.000
	Block	26.121	3	.000
	Model	26.121	3	.000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1017.787[a]	.033	.045

a. Estimation terminated at iteration number 3 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	.368	2	.832

Contingency Table for Hosmer and Lemeshow Test

		Home_Team_Win = 0		Home_Team_Win = 1		
		Observed	Expected	Observed	Expected	Total
Step 1	1	53	51.641	30	31.359	83
	2	16	17.359	21	19.641	37
	3	206	207.359	287	285.641	493
	4	46	44.641	109	110.359	155

Classification Table[a]

			Predicted		
			Home_Team_Win		Percentage Correct
	Observed		0	1	
Step 1	Home_Team_Win	0	57	264	17.8
		1	37	410	91.7
	Overall Percentage				60.8

a. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B)	
								Lower	Upper
Step 1[a]	H_CST	.585	.179	10.694	1	.001	1.795	1.264	2.548
	AO_EW	-.971	.311	9.740	1	.002	.379	.206	.697
	AR_F	-.701	.254	7.621	1	.006	.496	.302	.816
	Constant	.320	.089	12.852	1	.000	1.378		

a. Variable(s) entered on step 1: H_CST, AO_EW, AR_F.

Figure 3.5.
SPSS Results NFL 2016–2018 Away Team Win

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	25.833	3	.000
	Block	25.833	3	.000
	Model	25.833	3	.000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1015.341[a]	.033	.045

a. Estimation terminated at iteration number 3 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	.159	1	.690

Contingency Table for Hosmer and Lemeshow Test

		Away_Team_Win = 0		Away_Team_Win = 1		
		Observed	Expected	Observed	Expected	Total
Step 1	1	109	110.992	46	44.008	155
	2	290	288.008	203	204.992	493
	3	52	52.000	68	68.000	120

Classification Table[a]

Observed			Predicted Away_Team_Win 0	1	Percentage Correct
Step 1	Away_Team_Win	0	414	37	91.8
		1	260	57	18.0
	Overall Percentage				61.3

a. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B) Lower	Upper
Step 1[a]	H_CST	-.585	.179	10.640	1	.001	.557	.392	.792
	AO_EW	.991	.311	10.138	1	.001	2.694	1.464	4.959
	AR_F	.664	.253	6.858	1	.009	1.942	1.182	3.191
	Constant	-.340	.089	14.439	1	.000	.712		

a. Variable(s) entered on step 1: H_CST, AO_EW, AR_F.

MLS.

- DV is home team win. A LR was performed to ascertain the effects of Home team stadium surface type is field turf (HS_FT), and Away team time zone is mountain standard time (A_MST) on the likelihood if the Home_Team_Win. The LR model was statistically significant and all IVs were significant: $X^2(2) = 9.415$, $p = 0.009$. The model explained 0.8% (0.1%) of the variance in Home_Team_Win using Cox & Snell R^2 (Nagelkerke R^2) and correctly classified 53.2% of cases (see Figure 3.6).
- DV away team win. A LR was performed to ascertain the effects of Away team roof type is open (AR_O), and Away team roof type is retractable (AR_R) on the likelihood if the Away_Team_Win. The LR model was statistically significant and all IVs were significant: $X^2(1) = 9.372$, $p = 0.002$. The model explained .8% (1.3%) of the variance in the Away_Team_Win using Cox & Snell R^2 (Nagelkerke R^2) and correctly classified 78.6% of cases (see Figure 3.7).

Figure 3.6.
SPSS Results MLS 2016-2018 Home Team Win

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	9.415	2	.009
	Block	9.415	2	.009
	Model	9.415	2	.009

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1517.875[a]	.008	.011

a. Estimation terminated at iteration number 3 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	.120	1	.729

Contingency Table for Hosmer and Lemeshow Test

		Home_Team_Win = 0		Home_Team_Win = 1		
		Observed	Expected	Observed	Expected	Total
Step 1	1	396	398.156	409	406.844	805
	2	83	80.844	115	117.156	198
	3	38	38.000	64	64.000	102

Classification Table[a]

Observed			Predicted: Home_Team_Win 0	Predicted: Home_Team_Win 1	Percentage Correct
Step 1	Home_Team_Win	0	0	517	.0
		1	0	588	100.0
	Overall Percentage				53.2

a. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B) Lower	95% C.I.for EXP(B) Upper
Step 1[a]	HS_FT	.349	.153	5.195	1	.023	1.418	1.050	1.915
	A_MST	.424	.215	3.893	1	.048	1.528	1.003	2.328
	Constant	.022	.070	.095	1	.757	1.022		

a. Variable(s) entered on step 1: HS_FT, A_MST.

Figure 3.7.
SPSS Results MLS 2016-2018 Away Team Win

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	9.372	1	.002
	Block	9.372	1	.002
	Model	9.372	1	.002

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1139.456[a]	.008	.013

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	.000	0	.

Contingency Table for Hosmer and Lemeshow Test

		Away_Team_Win = 0		Away_Team_Win = 1		
		Observed	Expected	Observed	Expected	Total
Step 1	1	813	813.000	207	207.000	1020
	2	55	55.000	30	30.000	85

Classification Table[a]

			Predicted		
			Away_Team_Win		Percentage Correct
Observed			0	1	
Step 1	Away_Team_Win	0	868	0	100.0
		1	237	0	.0
	Overall Percentage				78.6

a. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B) Lower	95% C.I.for EXP(B) Upper
Step 1[a]	AR_O	-.762	.240	10.082	1	.001	.467	.292	.747
	Constant	-.606	.227	7.132	1	.008	.545		

a. Variable(s) entered on step 1: AR_O.

MLB.

- DV home team win. A LR was performed to ascertain the effects of (IVs) on the likelihood if the Home_Team_Win. The LR model was statistically significant, and all IVs were significant $X^2(25) = 50.839$, $p < 0.001$. The model explained 0.8% (1.1%) of the variance in the Home_Team_Win using Cox & Snell R^2 (Nagelkerke R^2) and correctly classified 53.6% of cases (see Figure 3.8).
- DV away team win. A LR was performed to ascertain the effects of (IVs) on the likelihood if the Away_Team_Win. The LR model was statistically significant, and all IVs were significant $X^2(25) = 59.461$, $p < 0.001$. The model explained 0.8 (1.1%) of the variance in the Away_Team_Win using Cox & Snell R^2 (Nagelkerke R^2) and correctly classified 53.6% of cases (see Figure 3.9).

Figure 3.8.

SPSS Results MLB 2016–2018 Home Team Win

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	59.839	25	.000
	Block	59.839	25	.000
	Model	59.839	25	.000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	10051.476[a]	.008	.011

a. Estimation terminated at iteration number 3 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	5.246	8	.731

Contingency Table for Hosmer and Lemeshow Test

		Home_ Team_Win = 0		Home_ Team_Win = 1		
		Observed	Expected	Observed	Expected	Total
Step 1	1	394	401.468	311	303.532	705
	2	416	408.297	331	338.703	747
	3	401	390.069	331	341.931	732
	4	358	378.075	373	352.925	731
	5	370	366.947	357	360.053	727
	6	363	358.323	365	369.677	728
	7	363	351.993	373	384.007	736
	8	326	338.320	402	389.680	728
	9	329	325.575	400	403.425	729
	10	303	303.933	428	427.067	731

Classification Table[a]

			Predicted		
			Home_ Team_Win		Percentage Correct
	Observed		0	1	
Step 1	Home_ Team_Win	0	1894	1729	52.3
		1	1657	2014	54.9
	Overall Percentage				53.6

a. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B) Lower	Upper
Step 1[a]	Travel_WE	.311	.138	5.068	1	.024	1.365	1.041	1.789
	Travel_EW	.283	.141	4.024	1	.045	1.327	1.007	1.750
	A_TZ_P3	.576	.681	.717	1	.397	1.779	.469	6.755
	A_TZ_P2	.495	.669	.548	1	.459	1.640	.442	6.083
	A_TZ_P1	.368	.453	.661	1	.416	1.445	.595	3.511
	A_TZ_0	.594	.359	2.746	1	.098	1.812	.897	3.660
	A_TZ_N1	.380	.309	1.507	1	.220	1.462	.797	2.680
	A_TZ_N2	.117	.156	.566	1	.452	1.124	.829	1.525
	H_PST	-.361	.317	1.304	1	.253	.697	.375	1.295
	H_MST	-.088	.297	.088	1	.767	.916	.512	1.638
	H_CST	.026	.100	.067	1	.796	1.026	.844	1.247
	H_Roof_Type_O	-.201	.085	5.632	1	.018	.818	.693	.966
	H_Roof_Type_F	-.122	.112	1.204	1	.273	.885	.711	1.101
	H_Playing_Surface_G	.235	.127	3.405	1	.065	1.265	.986	1.623
	HO_NWSE	-.099	.082	1.459	1	.227	.906	.772	1.063
	HO_SN	.032	.084	.145	1	.703	1.033	.876	1.217
	HO_SWNE	.141	.071	3.968	1	.046	1.151	1.002	1.322
	A_PDT	.306	.303	1.018	1	.313	1.357	.750	2.458
	A_MDT	.173	.276	.393	1	.531	1.189	.692	2.043
	A_Roof_Type_O	.167	.085	3.886	1	.049	1.182	1.001	1.396
	A_Roof_Type_F	.079	.111	.500	1	.479	1.082	.870	1.346
	A_Playing_Surface_G	-.235	.127	3.390	1	.066	.791	.616	1.015
	AO_NWSE	.091	.082	1.235	1	.267	1.095	.933	1.287
	AO_SN	-.122	.084	2.100	1	.147	.885	.751	1.044
	AO_SWNE	-.194	.071	7.409	1	.006	.824	.716	.947
	Constant	-.544	.409	1.769	1	.184	.581		

a. Variable(s) entered on step 1: Travel_WE, Travel_EW, A_TZ_P3, A_TZ_P2, A_TZ_P1, A_TZ_0, A_TZ_N1, A_TZ_N2, H_PST, H_MST, H_CST, H_Roof_Type_O, H_Roof_Type_F, H_Playing_Surface_G, HO_NWSE, HO_SN, HO_SWNE, A_PDT, A_MDT, A_Roof_Type_O, A_Roof_Type_F, A_Playing_Surface_G, AO_NWSE, AO_SN, AO_SWNE.

Figure 3.9.

SPSS Results MLB 2016–2018 Away Team Win

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	59.461	25	.000
	Block	59.461	25	.000
	Model	59.461	25	.000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	10051.828[a]	.008	.011

a. Estimation terminated at iteration number 3 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	4.914	8	.767

Contingency Table for Hosmer and Lemeshow Test

		Away_Team_Win = 0		Away_Team_Win = 1		
		Observed	Expected	Observed	Expected	Total
Step 1	1	428	426.967	303	304.033	731
	2	405	406.679	330	328.321	735
	3	397	389.583	331	338.417	728
	4	373	380.947	357	349.053	730
	5	371	375.763	369	364.237	740
	6	355	360.423	373	367.577	728
	7	369	346.797	349	371.203	718
	8	329	338.361	395	385.639	724
	9	334	342.789	421	412.211	755
	10	311	303.691	394	401.309	705

Classification Table[a]

			Predicted		
			Away_Team_Win		Percentage Correct
Observed			0	1	
Step 1	Away_Team_Win	0	2014	1658	54.8
		1	1729	1893	52.3
	Overall Percentage				53.6

a. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B) Lower	Upper
Step 1[a]	Travel_WE	-.311	.138	5.082	1	.024	.732	.559	.960
	Travel_EW	-.283	.141	4.024	1	.045	.753	.571	.994
	A_TZ_P3	-.570	.681	.702	1	.402	.565	.149	2.146
	A_TZ_P2	-.490	.669	.537	1	.464	.613	.165	2.272
	A_TZ_P1	-.366	.453	.652	1	.419	.694	.286	1.685
	A_TZ_0	-.592	.359	2.722	1	.099	.553	.274	1.118
	A_TZ_N1	-.380	.309	1.513	1	.219	.684	.373	1.253
	A_TZ_N2	-.117	.156	.565	1	.452	.890	.656	1.207
	H_PST	.358	.316	1.280	1	.258	1.431	.769	2.660
	H_MST	.087	.297	.086	1	.769	1.091	.610	1.951
	H_CST	-.025	.100	.064	1	.801	.975	.802	1.186
	H_Roof_Type_O	.201	.085	5.625	1	.018	1.223	1.036	1.444
	H_Roof_Type_F	.123	.112	1.216	1	.270	1.131	.909	1.407
	H_Playing_Surface_G	-.235	.127	3.420	1	.064	.790	.616	1.014
	HO_NWSE	.096	.082	1.377	1	.241	1.101	.938	1.292
	HO_SN	-.032	.084	.146	1	.703	.968	.821	1.142
	HO_SWNE	-.141	.071	3.985	1	.046	.868	.756	.997
	A_PDT	-.302	.303	.995	1	.319	.739	.408	1.339
	A_MDT	-.170	.276	.378	1	.539	.844	.491	1.450
	A_Roof_Type_O	-.168	.085	3.926	1	.048	.845	.716	.998
	A_Roof_Type_F	-.079	.111	.499	1	.480	.924	.743	1.150
	A_Playing_Surface_G	.235	.127	3.402	1	.065	1.265	.985	1.624
	AO_NWSE	-.091	.082	1.220	1	.269	.913	.778	1.073
	AO_SN	.121	.084	2.087	1	.149	1.129	.958	1.331
	AO_SWNE	.193	.071	7.313	1	.007	1.212	1.054	1.394
	Constant	.543	.409	1.763	1	.184	1.721		

a. Variable(s) entered on step 1: Travel_WE, Travel_EW, A_TZ_P3, A_TZ_P2, A_TZ_P1, A_TZ_0, A_TZ_N1, A_TZ_N2, H_PST, H_MST, H_CST, H_Roof_Type_O, H_Roof_Type_F, H_Playing_Surface_G, HO_NWSE, HO_SN, HO_SWNE, A_PDT, A_MDT, A_Roof_Type_O, A_Roof_Type_F, A_Playing_Surface_G, AO_NWSE, AO_SN, AO_SWNE.

Two other scenarios for each sport were tested in addition to all previous analysis scenarios.

1. One single season analysis
    1.1. Single season of NFL (2016)
    1.2. Single season of MLS (2016)
    1.3. Single season of MLB (2016)
2. One team analysis
    2.1. One team best record of the competition
        2.1.1. One team best record of NFL (2018, Rams)
        2.1.2. One team best record of MLS (2018, NY Red Bulls)
        2.1.3. One team best record of MLB (2018, Red Sox)
    2.2. One team worst record
        2.2.1. One team worst record of NFL (2018, Cardinals)
        2.2.2. One team worst record of MLS (2018, San Jose)
        2.2.3. One team worst record of MLB (2018, Orioles)
    2.3. One team (middle-range)
        2.3.1. One team average record of NFL (2018, Packers)
        2.3.2. One team average record of MLS (2018, LA Galaxy)
        2.3.3. One team average record of MLB (2018, Diamondbacks)

The results in some scenarios were similar to the previous findings; however, some results were different or not statistically significant to the findings previously presented. See Appendix C for all tests results. All results from SPSS present tables of statistics. The following are short briefs of each table.

The Omnibus tests of model coefficients. According to ReStore: National Centre for Research Methods (2011), the omnibus tests of model coefficients is "used to check that the new model—with explanatory variables included—is an improvement over the baseline model."

Hosmer and Lemeshow test. This table shows two results: Hosmer-Lemeshow chi-squared and a *p*-value. The *p*-value indicates if the model is a good or poor fit (Stephanie, 2016).

Classification table. The model's classificatory power—how well the model predicts the two outcomes—is called the prediction success table. This model is another way of evaluating the fit of a given LR model.

Variables in the equation table. This type of table shows the contribution of each independent variable to the model and its statistical significance. The Wald test is used to determine statistical significance for each of the independent variables. The statistical significance of the test is found in the Sig column.

Conclusion

A LR analysis of the NFL, MLS, and MLB indicated that a significant relationship exists between spatial, temporal, and stadium attributes and the game results. The results vary from one sport to another.

Three IVs were statistically significant in the NFL: a) when the home team is from the central standard time zone (H_CST), b) when the away team stadium orientation is east to west (AO_EW), and c) when away team's stadium roof type is fixed (AR_F).

Several IVs were correlated for MLS, particularly when the DV is the home team winning. Two significant variables were discovered when a) the home team stadium surface is field turf (HS_FT), and b) the away team is from the Mountain Standard Time zone (A_MST).

Two IVs were statistically significant when the DV is the away team winning: a) when the away team stadium roof type is open (AR_O) and b) when the stadium roof is retractable (AR_R).

For MLB, six IVs were statistically significant for both cases when the DV is the home team winning or away team winning. The first IV was when the away team stadium orientation is southwest to northeast (AO_SWNE). The second was when the home team stadium roof type is open (HR_O), followed by the away team traveling from west to east (Travel_WE), the away team traveling from east to west (Travel_EW), and a home team stadium orientation of southwest to northeast (HO_SWNE), The last IV was when the away team stadium roof type is open (AR_O).

References

Cabrera, A. F. (1994). Logistic regression analysis in higher education: An applied perspective. In J.C. Smart (Ed.), *Higher education: Handbook of theory and research* (pp. 225–256).

Courneya, K. S., & Carron, A. V. (1992). The home advantage in sport competitions: A literature review. *Journal of Sport and Exercise Psychology, 14*(1), 13–27. https://doi.org/10.1123/jsep.12.3.312

Harper, G., & Pickett, S. D. (2006). Methods for mining HTS data. *Drug Discovery Today, 11*(15–16), 694–699. https://doi.org/10.1016/j.drudis.2006.06.006

IBM Knowledge Center. (2012). *Parameter settings.* https://www.ibm.com/support/knowledgecenter/SS3RA7_15.0.0/com.ibm.spss.crispdm.help/crisp_parameter_settings.htm

Laerd Statistics. (n.d.). *Binomial logistic regression using SPSS statistics.* (n.d.) https://statistics.laerd.com/spss-tutorials/binomial-logistic-regression-using-spss-statistics.php#procedure

Locke, S. (2017, January 13). *CRISP-DM and why you should know about it.* Locke Data. https://itsalocke.com/blog/crisp-dm-and-why-you-should-know-about-it/

Peng, C. Y. J., Lee, K. L., & Ingersoll, G. M. (2002). An introduction to logistic regression analysis and reporting. *The Journal of Educational Research, 96*(1), 3–14. https://doi.org/10.1080/00220670209598786

ReStore: National Centre for Research Methods. (2011, July 22). *The SPSS logistic regression output.* http://www.restore.ac.uk/srme/www/fac/soc/wie/research-new/srme/modules/mod4/12/index.html

Smart Vision. (n.d.). *What is the CRISP-DM methodology?* https://www.sv-europe.com/crisp-dm-methodology/

Wirth, R., & Hipp, J. (2000). *CRISP-DM: Towards a standard process model for data mining* [Paper presentation]. Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining.

Appendix A: Stadium Playing Surface and Stadium Roof Type References

Kanses City Chiefs. (2016). *Chiefs 2016 media guide*. http://prod.static.chiefs.clubs.nfl.com/assets/pdf/2016/2016MediaGuide.pdf#page=382

National Football League. (2016). *Official 2016 National Football League record and fact book.* https://nflcommunications.com/Documents/2016%20NFL%20Record%20and%20Fact%20Book.pdf#page=161

Dallas Cowboys. (2017). *2017 media guide.* http://prod.static.cowboys.clubs.nfl.com/assets/docs/Dallas-Cowboys-Media-Guide.pdf#page=3

Carolina Panthers. (n.d.). *Stadium facts.* https://www.panthers.com/stadium/facts

Empower Field at Mile High. (n.d.). *Facts and figures.* https://www.empowerfieldatmilehigh.com/stadium-information/about-us/facts-figures

Century Link Field. (n.d.). *Stadium facts.* http://www.centurylinkfield.com/stadium-facts/

The City of San Diego. (n.d.). *Stadium fact guide.* https://www.sandiego.gov/stadium/about/factguide

First Energy Stadium. (n.d.). *About us.* http://firstenergystadium.com/stadium-info/about-us/

Washington Redskins. (2017). *2017 Washington Redskins media guide.* http://prod.static.redskins.clubs.nfl.com/assets/docs/Washington-Redskins-Media-Guide.pdf#page=380

Detroit Lions. (n.d.). *Ford Field facts & history.* https://web.archive.org/web/20150905082512/http:/www.detroitlions.com/ford-field/facts-history.html

Gillette Stadium. (n.d.). *Overview.* https://www.gillettestadium.com/overview-venue/#quickhits

Heinz Field. (n.d.). *Heinz Field facts.* https://heinzfield.com/stadium/heinz-field-facts/

Levi’s Stadium. (n.d.). *About.* https://www.levisstadium.com/about-levis-stadium/

Tennessee Titans. (2016). *Tennessee Titans 2016 media guide.* http://prod.static.titans.clubs.nfl.com/assets/docs/mediaguide/2016TennesseeTitansMediaGuide.pdf#page=2

Oakland Raiders. (2017). *Raiders 2017.* http://prod.static.raiders.clubs.nfl.com/assets/docs/Oakland-Raiders-Media-Guide.pdf#page=9

University of Phoenix Stadium. (n.d.). *Stadium history.* https://web.archive.org/web/20150820102606/http:/www.universityofphoenixstadium.com/stadium

Silverstein, T. (2012, January 14). *Lambeau surface kept safe and soft through technology.* Journal Sentinel. http://archive.jsonline.com/sports/packers/lambeau-surface-kept-safe-and-soft-through-technology-763p13j-137327223.html/

NBC Sports Philadelphia. (2013, October 24). *Linc's structure, playing surface a bigger problem than family-friendly atmosphere.* https://www.nbcsports.com/philadelphia/lincs-structure-playing-surface-bigger-problem-family-friendly-atmosphere

Atlanta Falcons. (2017, April 4). *Mercedes-Benz stadium will have field turf.* https://www.atlantafalcons.com/news/mercedes-benz-stadium-will-have-fieldturf-18709324

O'Halloran, R. (2018, February 12). *Jaguars announce tarp removal , 2018 season-ticket renewal plan.* The Florida Times Union. https://www.jacksonville.com/sports/20180212/jaguars-announce-tarp-removal-2018-season-ticket-renewal-plan

Appendix B: Correlation Tests for all sports

MLS Correlation Between Variables For EC_Home Team Wins, Away Team Wins, And Stadium Orientation

Correlations

		Home_ Team_Win	Away_Team_ Win	Draw	AO_NESW	AO_NS	AO_EW
Home_ Team_Win	Pearson Correlation	1	-.579**	-.609**	-.026	.287*	-.346**
	Sig. (2-tailed)		.000	.000	.853	.034	.010
	N	55	55	55	55	55	55
Away_Team_Win	Pearson Correlation	-.579**	1	-.294*	-.148	-.123	.292*
	Sig. (2-tailed)	.000		.029	.281	.370	.030
	N	55	55	55	55	55	55
Draw	Pearson Correlation	-.609**	-.294*	1	.174	-.217	.122
	Sig. (2-tailed)	.000	.029		.204	.112	.376
	N	55	55	55	55	55	55
AO_NESW	Pearson Correlation	-.026	-.148	.174	1	-.633**	-.089
	Sig. (2-tailed)	.853	.281	.204		.000	.520
	N	55	55	55	55	55	55
AO_NS	Pearson Correlation	.287*	-.123	-.217	-.633**	1	-.715**
	Sig. (2-tailed)	.034	.370	.112	.000		.000
	N	55	55	55	55	55	55
AO_EW	Pearson Correlation	-.346**	.292*	.122	-.089	-.715**	1
	Sig. (2-tailed)	.010	.030	.376	.520	.000	
	N	55	55	55	55	55	55

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Correlation Between Variables For WC Home Team Wins, Away Team Wins, And Stadium Orientation

Correlations

		Home_ Team_Win	Away_Team_ Win	Draw	AO_NESW	AO_NS	AO_EW
Home_ Team_Win	Pearson Correlation	1	-.548**	-.600**	-.226	.016	.342**
	Sig. (2-tailed)		.000	.000	.086	.904	.008
	N	59	59	59	59	59	59
Away_Team_Win	Pearson Correlation	-.548**	1	-.340**	.086	.028	-.188
	Sig. (2-tailed)	.000		.008	.517	.831	.155
	N	59	59	59	59	59	59
Draw	Pearson Correlation	-.600**	-.340**	1	.171	-.045	-.205
	Sig. (2-tailed)	.000	.008		.195	.733	.119
	N	59	59	59	59	59	59
AO_NESW	Pearson Correlation	-.226	.086	.171	1	-.815**	-.289*
	Sig. (2-tailed)	.086	.517	.195		.000	.027
	N	59	59	59	59	59	59
AO_NS	Pearson Correlation	.016	.028	-.045	-.815**	1	-.320*
	Sig. (2-tailed)	.904	.831	.733	.000		.014
	N	59	59	59	59	59	59
AO_EW	Pearson Correlation	.342**	-.188	-.205	-.289*	-.320*	1
	Sig. (2-tailed)	.008	.155	.119	.027	.014	
	N	59	59	59	59	59	59

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Correlation Tests for NFL

Correlation Between Variables For All NFL Away Team Wins And Away Time Zone Differences

Correlations

		Away_ Team_Win	A_TZ_P3	A_TZ_P2	A_TZ_P1	A_TZ_0	A_TZ_N1	A_TZ_N2	A_TZ_N3
Away_ Team_Win	Pearson Correlation	1	.039	.003	-.115	-.061	.187**	-.103	.085
	Sig. (2-tailed)		.537	.965	.065	.330	.003	.099	.176
	N	256	256	256	256	256	256	256	256
A_TZ_P3	Pearson Correlation	.039	1	-.045	-.088	-.157*	-.088	-.040	-.036
	Sig. (2-tailed)	.537		.475	.158	.012	.158	.526	.564
	N	256	256	256	256	256	256	256	256
A_TZ_P2	Pearson Correlation	.003	-.045	1	-.123*	-.218**	-.123*	-.055	-.050
	Sig. (2-tailed)	.965	.475		.049	.000	.049	.378	.423
	N	256	256	256	256	256	256	256	256
A_TZ_P1	Pearson Correlation	-.115	-.088	-.123*	1	-.431**	-.243**	-.109	-.099
	Sig. (2-tailed)	.065	.158	.049		.000	.000	.081	.113
	N	256	256	256	256	256	256	256	256
A_TZ_0	Pearson Correlation	-.061	-.157*	-.218**	-.431**	1	-.431**	-.194**	-.176**
	Sig. (2-tailed)	.330	.012	.000	.000		.000	.002	.005
	N	256	256	256	256	256	256	256	256
A_TZ_N1	Pearson Correlation	.187**	-.088	-.123*	-.243**	-.431**	1	-.109	-.099
	Sig. (2-tailed)	.003	.158	.049	.000	.000		.081	.113
	N	256	256	256	256	256	256	256	256
A_TZ_N2	Pearson Correlation	-.103	-.040	-.055	-.109	-.194**	-.109	1	-.045
	Sig. (2-tailed)	.099	.526	.378	.081	.002	.081		.476
	N	256	256	256	256	256	256	256	256
A_TZ_N3	Pearson Correlation	.085	-.036	-.050	-.099	-.176**	-.099	-.045	1
	Sig. (2-tailed)	.176	.564	.423	.113	.005	.113	.476	
	N	256	256	256	256	256	256	256	256

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Home Team Wins and Away Time Zone Differences

Correlations

		Home_Team_Win	A_TZ_P3	A_TZ_P2	A_TZ_P1	A_TZ_0	A_TZ_N1	A_TZ_N2	A_TZ_N3
Home_Team_Win	Pearson Correlation	1	-.036	.001	.123*	.043	-.178**	.107	-.081
	Sig. (2-tailed)		.569	.985	.049	.496	.004	.089	.195
	N	256	256	256	256	256	256	256	256
A_TZ_P3	Pearson Correlation	-.036	1	-.045	-.088	-.157*	-.088	-.040	-.036
	Sig. (2-tailed)	.569		.475	.158	.012	.158	.526	.564
	N	256	256	256	256	256	256	256	256
A_TZ_P2	Pearson Correlation	.001	-.045	1	-.123*	-.218**	-.123*	-.055	-.050
	Sig. (2-tailed)	.985	.475		.049	.000	.049	.378	.423
	N	256	256	256	256	256	256	256	256
A_TZ_P1	Pearson Correlation	.123*	-.088	-.123*	1	-.431**	-.243**	-.109	-.099
	Sig. (2-tailed)	.049	.158	.049		.000	.000	.081	.113
	N	256	256	256	256	256	256	256	256
A_TZ_0	Pearson Correlation	.043	-.157*	-.218**	-.431**	1	-.431**	-.194**	-.176**
	Sig. (2-tailed)	.496	.012	.000	.000		.000	.002	.005
	N	256	256	256	256	256	256	256	256
A_TZ_N1	Pearson Correlation	-.178**	-.088	-.123*	-.243**	-.431**	1	-.109	-.099
	Sig. (2-tailed)	.004	.158	.049	.000	.000		.081	.113
	N	256	256	256	256	256	256	256	256
A_TZ_N2	Pearson Correlation	.107	-.040	-.055	-.109	-.194**	-.109	1	-.045
	Sig. (2-tailed)	.089	.526	.378	.081	.002	.081		.476
	N	256	256	256	256	256	256	256	256
A_TZ_N3	Pearson Correlation	-.081	-.036	-.050	-.099	-.176**	-.099	-.045	1
	Sig. (2-tailed)	.195	.564	.423	.113	.005	.113	.476	
	N	256	256	256	256	256	256	256	256

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Away Team Wins and Away Team Stadium Orientation

Correlations

		Away_ Team_Win	AO_EW	AO_NESW	AO_NS	AO_NWSE
Away_ Team_Win	Pearson Correlation	1	.089	-.039	.153*	-.155*
	Sig. (2-tailed)		.157	.532	.014	.013
	N	256	256	256	256	256
AO_EW	Pearson Correlation	.089	1	-.098	-.162**	-.275**
	Sig. (2-tailed)	.157		.119	.010	.000
	N	256	256	256	256	256
AO_NESW	Pearson Correlation	-.039	-.098	1	-.236**	-.402**
	Sig. (2-tailed)	.532	.119		.000	.000
	N	256	256	256	256	256
AO_NS	Pearson Correlation	.153*	-.162**	-.236**	1	-.666**
	Sig. (2-tailed)	.014	.010	.000		.000
	N	256	256	256	256	256
AO_NWSE	Pearson Correlation	-.155*	-.275**	-.402**	-.666**	1
	Sig. (2-tailed)	.013	.000	.000	.000	
	N	256	256	256	256	256

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Away Team Wins and Home Team Stadium Orientation

Correlations

		Away_ Team_Win	HO_EW	HO_NESW	HO_NS	HO_NWSE
Away_ Team_Win	Pearson Correlation	1	-.043	-.063	-.078	.134*
	Sig. (2-tailed)		.490	.312	.212	.033
	N	256	256	256	256	256
HO_EW	Pearson Correlation	-.043	1	-.098	-.162**	-.275**
	Sig. (2-tailed)	.490		.119	.010	.000
	N	256	256	256	256	256
HO_NESW	Pearson Correlation	-.063	-.098	1	-.236**	-.402**
	Sig. (2-tailed)	.312	.119		.000	.000
	N	256	256	256	256	256
HO_NS	Pearson Correlation	-.078	-.162**	-.236**	1	-.666**
	Sig. (2-tailed)	.212	.010	.000		.000
	N	256	256	256	256	256
HO_NWSE	Pearson Correlation	.134*	-.275**	-.402**	-.666**	1
	Sig. (2-tailed)	.033	.000	.000	.000	
	N	256	256	256	256	256

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Home Team Wins and Away Team Stadium Orientation

Correlations

		Home_Team_Win	AO_EW	AO_NESW	AO_NS	AO_NWSE
Home_Team_Win	Pearson Correlation	1	-.084	.045	-.142*	.139*
	Sig. (2-tailed)		.179	.472	.023	.026
	N	256	256	256	256	256
AO_EW	Pearson Correlation	-.084	1	-.098	-.162**	-.275**
	Sig. (2-tailed)	.179		.119	.010	.000
	N	256	256	256	256	256
AO_NESW	Pearson Correlation	.045	-.098	1	-.236**	-.402**
	Sig. (2-tailed)	.472	.119		.000	.000
	N	256	256	256	256	256
AO_NS	Pearson Correlation	-.142*	-.162**	-.236**	1	-.666**
	Sig. (2-tailed)	.023	.010	.000		.000
	N	256	256	256	256	256
AO_NWSE	Pearson Correlation	.139*	-.275**	-.402**	-.666**	1
	Sig. (2-tailed)	.026	.000	.000	.000	
	N	256	256	256	256	256

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Away Team Wins and Travel Direction

Correlations

		Away_ Team_Win	Travel_WE	Travel_EW
Away_ Team_Win	Pearson Correlation	1	.153*	-.085
	Sig. (2-tailed)		.014	.175
	N	256	256	256
Travel_WE	Pearson Correlation	.153*	1	-.395**
	Sig. (2-tailed)	.014		.000
	N	256	256	256
Travel_EW	Pearson Correlation	-.085	-.395**	1
	Sig. (2-tailed)	.175	.000	
	N	256	256	256

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Home Team Wins and Travel Direction

Correlations

		Home_Team_Win	Travel_WE	Travel_EW
Home_Team_Win	Pearson Correlation	1	-.142*	.095
	Sig. (2-tailed)		.023	.130
	N	256	256	256
Travel_WE	Pearson Correlation	-.142*	1	-.395**
	Sig. (2-tailed)	.023		.000
	N	256	256	256
Travel_EW	Pearson Correlation	.095	-.395**	1
	Sig. (2-tailed)	.130	.000	
	N	256	256	256

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Correlation for American Conference in NFL

Away Team Win and Away Team Stadium Orientation (AO_NS)

Correlations

		Away_ Team_Win	AO_EW	AO_NESW	AO_NS	AO_NWSE
Away_ Team_Win	Pearson Correlation	1	.[a]	-.041	.289**	-.200
	Sig. (2-tailed)		.	.689	.004	.051
	N	96	96	96	96	96
AO_EW	Pearson Correlation	.[a]	.[a]	.[a]	.[a]	.[a]
	Sig. (2-tailed)	.		.	.	.
	N	96	96	96	96	96
AO_NESW	Pearson Correlation	-.041	.[a]	1	-.231*	-.620**
	Sig. (2-tailed)	.689	.		.024	.000
	N	96	96	96	96	96
AO_NS	Pearson Correlation	.289**	.[a]	-.231*	1	-.620**
	Sig. (2-tailed)	.004	.	.024		.000
	N	96	96	96	96	96
AO_NWSE	Pearson Correlation	-.200	.[a]	-.620**	-.620**	1
	Sig. (2-tailed)	.051	.	.000	.000	
	N	96	96	96	96	96

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

a. Cannot be computed because at least one of the variables is constant.

Home Team Win and Away Team Stadium Orientation

Correlations

		Home_Team_Win	AO_EW	AO_NESW	AO_NS	AO_NWSE
Home_Team_Win	Pearson Correlation	1	.[a]	.051	-.278**	.182
	Sig. (2-tailed)		.	.619	.006	.075
	N	96	96	96	96	96
AO_EW	Pearson Correlation	.[a]	.[a]	.[a]	.[a]	.[a]
	Sig. (2-tailed)	.		.	.	.
	N	96	96	96	96	96
AO_NESW	Pearson Correlation	.051	.[a]	1	-.231*	-.620**
	Sig. (2-tailed)	.619	.		.024	.000
	N	96	96	96	96	96
AO_NS	Pearson Correlation	-.278**	.[a]	-.231*	1	-.620**
	Sig. (2-tailed)	.006	.	.024		.000
	N	96	96	96	96	96
AO_NWSE	Pearson Correlation	.182	.[a]	-.620**	-.620**	1
	Sig. (2-tailed)	.075	.	.000	.000	
	N	96	96	96	96	96

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

a. Cannot be computed because at least one of the variables is constant.

Correlation for National Conference in NFL

Away Team Win and Away Time Zone Differences

Correlations

		Away_ Team_Win	A_TZ_P3	A_TZ_P2	A_TZ_P1	A_TZ_0	A_TZ_N1	A_TZ_N2	A_TZ_N3
Away_ Team_Win	Pearson Correlation	1	.095	-.019	-.145	-.071	.255*	-.214*	.146
	Sig. (2-tailed)		.356	.853	.159	.491	.012	.037	.156
	N	96	96	96	96	96	96	96	96
A_TZ_P3	Pearson Correlation	.095	1	-.054	-.089	-.139	-.092	-.046	-.037
	Sig. (2-tailed)	.356		.600	.387	.176	.372	.654	.717
	N	96	96	96	96	96	96	96	96
A_TZ_P2	Pearson Correlation	-.019	-.054	1	-.150	-.234*	-.155	-.078	-.063
	Sig. (2-tailed)	.853	.600		.145	.022	.132	.451	.543
	N	96	96	96	96	96	96	96	96
A_TZ_P1	Pearson Correlation	-.145	-.089	-.150	1	-.385**	-.255*	-.128	-.104
	Sig. (2-tailed)	.159	.387	.145		.000	.012	.213	.315
	N	96	96	96	96	96	96	96	96
A_TZ_0	Pearson Correlation	-.071	-.139	-.234*	-.385**	1	-.397**	-.200	-.162
	Sig. (2-tailed)	.491	.176	.022	.000		.000	.051	.116
	N	96	96	96	96	96	96	96	96
A_TZ_N1	Pearson Correlation	.255*	-.092	-.155	-.255*	-.397**	1	-.132	-.107
	Sig. (2-tailed)	.012	.372	.132	.012	.000		.198	.300
	N	96	96	96	96	96	96	96	96
A_TZ_N2	Pearson Correlation	-.214*	-.046	-.078	-.128	-.200	-.132	1	-.054
	Sig. (2-tailed)	.037	.654	.451	.213	.051	.198		.602
	N	96	96	96	96	96	96	96	96
A_TZ_N3	Pearson Correlation	.146	-.037	-.063	-.104	-.162	-.107	-.054	1
	Sig. (2-tailed)	.156	.717	.543	.315	.116	.300	.602	
	N	96	96	96	96	96	96	96	96

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Home Team Win and Away Time zone Difference:

Correlations

		Home_Team_Win	A_TZ_P3	A_TZ_P2	A_TZ_P1	A_TZ_0	A_TZ_N1	A_TZ_N2	A_TZ_N3
Home_Team_Win	Pearson Correlation	1	-.091	.025	.155	.044	-.243*	.218*	-.141
	Sig. (2-tailed)		.378	.805	.132	.673	.017	.033	.171
	N	96	96	96	96	96	96	96	96
A_TZ_P3	Pearson Correlation	-.091	1	-.054	-.089	-.139	-.092	-.046	-.037
	Sig. (2-tailed)	.378		.600	.387	.176	.372	.654	.717
	N	96	96	96	96	96	96	96	96
A_TZ_P2	Pearson Correlation	.025	-.054	1	-.150	-.234*	-.155	-.078	-.063
	Sig. (2-tailed)	.805	.600		.145	.022	.132	.451	.543
	N	96	96	96	96	96	96	96	96
A_TZ_P1	Pearson Correlation	.155	-.089	-.150	1	-.385**	-.255*	-.128	-.104
	Sig. (2-tailed)	.132	.387	.145		.000	.012	.213	.315
	N	96	96	96	96	96	96	96	96
A_TZ_0	Pearson Correlation	.044	-.139	-.234*	-.385**	1	-.397**	-.200	-.162
	Sig. (2-tailed)	.673	.176	.022	.000		.000	.051	.116
	N	96	96	96	96	96	96	96	96
A_TZ_N1	Pearson Correlation	-.243*	-.092	-.155	-.255*	-.397**	1	-.132	-.107
	Sig. (2-tailed)	.017	.372	.132	.012	.000		.198	.300
	N	96	96	96	96	96	96	96	96
A_TZ_N2	Pearson Correlation	.218*	-.046	-.078	-.128	-.200	-.132	1	-.054
	Sig. (2-tailed)	.033	.654	.451	.213	.051	.198		.602
	N	96	96	96	96	96	96	96	96
A_TZ_N3	Pearson Correlation	-.141	-.037	-.063	-.104	-.162	-.107	-.054	1
	Sig. (2-tailed)	.171	.717	.543	.315	.116	.300	.602	
	N	96	96	96	96	96	96	96	96

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Away Team Wins and Home Team Stadium Orientation

Correlations

		Away_ Team_Win	HO_EW	HO_NESW	HO_NS	HO_NWSE
Away_ Team_Win	Pearson Correlation	1	-.120	-.214*	-.159	.339**
	Sig. (2-tailed)		.243	.037	.122	.001
	N	96	96	96	96	96
HO_EW	Pearson Correlation	-.120	1	-.098	-.293**	-.333**
	Sig. (2-tailed)	.243		.344	.004	.001
	N	96	96	96	96	96
HO_NESW	Pearson Correlation	-.214*	-.098	1	-.200	-.228*
	Sig. (2-tailed)	.037	.344		.051	.026
	N	96	96	96	96	96
HO_NS	Pearson Correlation	-.159	-.293**	-.200	1	-.683**
	Sig. (2-tailed)	.122	.004	.051		.000
	N	96	96	96	96	96
HO_NWSE	Pearson Correlation	.339**	-.333**	-.228*	-.683**	1
	Sig. (2-tailed)	.001	.001	.026	.000	
	N	96	96	96	96	96

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Home Team Wins and Home Team Stadium Orientation

Correlations

		Home_Team_Win	HO_EW	HO_NESW	HO_NS	HO_NWSE
Home_Team_Win	Pearson Correlation	1	.128	.218*	.131	-.319**
	Sig. (2-tailed)		.215	.033	.204	.002
	N	96	96	96	96	96
HO_EW	Pearson Correlation	.128	1	-.098	-.293**	-.333**
	Sig. (2-tailed)	.215		.344	.004	.001
	N	96	96	96	96	96
HO_NESW	Pearson Correlation	.218*	-.098	1	-.200	-.228*
	Sig. (2-tailed)	.033	.344		.051	.026
	N	96	96	96	96	96
HO_NS	Pearson Correlation	.131	-.293**	-.200	1	-.683**
	Sig. (2-tailed)	.204	.004	.051		.000
	N	96	96	96	96	96
HO_NWSE	Pearson Correlation	-.319**	-.333**	-.228*	-.683**	1
	Sig. (2-tailed)	.002	.001	.026	.000	
	N	96	96	96	96	96

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Correlation Tests for MLB

Away Team Wins and Away Time Zone Differences

Correlations

		Away_Team_Win	A_TZ_P3	A_TZ_P2	A_TZ_P1	A_TZ_0	A_TZ_N1	A_TZ_N2	A_TZ_N3
Away_Team_Win	Pearson Correlation	1	-.011	.003	.012	-.030	-.016	.049*	.025
	Sig. (2-tailed)		.581	.892	.567	.144	.418	.016	.216
	N	2434	2434	2434	2434	2434	2434	2434	2434
A_TZ_P3	Pearson Correlation	-.011	1	-.071**	-.104**	-.240**	-.106**	-.071**	-.070**
	Sig. (2-tailed)	.581		.000	.000	.000	.000	.000	.000
	N	2434	2487	2487	2487	2487	2487	2487	2487
A_TZ_P2	Pearson Correlation	.003	-.071**	1	-.107**	-.248**	-.109**	-.073**	-.072**
	Sig. (2-tailed)	.892	.000		.000	.000	.000	.000	.000
	N	2434	2487	2487	2487	2487	2487	2487	2487
A_TZ_P1	Pearson Correlation	.012	-.104**	-.107**	1	-.364**	-.161**	-.107**	-.106**
	Sig. (2-tailed)	.567	.000	.000		.000	.000	.000	.000
	N	2434	2487	2487	2487	2487	2487	2487	2487
A_TZ_0	Pearson Correlation	-.030	-.240**	-.248**	-.364**	1	-.371**	-.247**	-.244**
	Sig. (2-tailed)	.144	.000	.000	.000		.000	.000	.000
	N	2434	2487	2487	2487	2487	2487	2487	2487
A_TZ_N1	Pearson Correlation	-.016	-.106**	-.109**	-.161**	-.371**	1	-.109**	-.108**
	Sig. (2-tailed)	.418	.000	.000	.000	.000		.000	.000
	N	2434	2487	2487	2487	2487	2487	2487	2487
A_TZ_N2	Pearson Correlation	.049*	-.071**	-.073**	-.107**	-.247**	-.109**	1	-.072**
	Sig. (2-tailed)	.016	.000	.000	.000	.000	.000		.000
	N	2434	2487	2487	2487	2487	2487	2487	2487
A_TZ_N3	Pearson Correlation	.025	-.070**	-.072**	-.106**	-.244**	-.108**	-.072**	1
	Sig. (2-tailed)	.216	.000	.000	.000	.000	.000	.000	
	N	2434	2487	2487	2487	2487	2487	2487	2487

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Home Team Win and Away Time Zone Differences

Correlations

		Home_ Team_Win	A_TZ_P3	A_TZ_P2	A_TZ_P1	A_TZ_0	A_TZ_N1	A_TZ_N2	A_TZ_N3
Home_ Team_Win	Pearson Correlation	1	.009	-.006	.008	.028	.013	-.057**	-.031
	Sig. (2-tailed)		.650	.761	.704	.170	.511	.005	.132
	N	2432	2432	2432	2432	2432	2432	2432	2432
A_TZ_P3	Pearson Correlation	.009	1	-.071**	-.104**	-.240**	-.106**	-.071**	-.070**
	Sig. (2-tailed)	.650		.000	.000	.000	.000	.000	.000
	N	2432	2487	2487	2487	2487	2487	2487	2487
A_TZ_P2	Pearson Correlation	-.006	-.071**	1	-.107**	-.248**	-.109**	-.073**	-.072**
	Sig. (2-tailed)	.761	.000		.000	.000	.000	.000	.000
	N	2432	2487	2487	2487	2487	2487	2487	2487
A_TZ_P1	Pearson Correlation	.008	-.104**	-.107**	1	-.364**	-.161**	-.107**	-.106**
	Sig. (2-tailed)	.704	.000	.000		.000	.000	.000	.000
	N	2432	2487	2487	2487	2487	2487	2487	2487
A_TZ_0	Pearson Correlation	.028	-.240**	-.248**	-.364**	1	-.371**	-.247**	-.244**
	Sig. (2-tailed)	.170	.000	.000	.000		.000	.000	.000
	N	2432	2487	2487	2487	2487	2487	2487	2487
A_TZ_N1	Pearson Correlation	.013	-.106**	-.109**	-.161**	-.371**	1	-.109**	-.108**
	Sig. (2-tailed)	.511	.000	.000	.000	.000		.000	.000
	N	2432	2487	2487	2487	2487	2487	2487	2487
A_TZ_N2	Pearson Correlation	-.057**	-.071**	-.073**	-.107**	-.247**	-.109**	1	-.072**
	Sig. (2-tailed)	.005	.000	.000	.000	.000	.000		.000
	N	2432	2487	2487	2487	2487	2487	2487	2487
A_TZ_N3	Pearson Correlation	-.031	-.070**	-.072**	-.106**	-.244**	-.108**	-.072**	1
	Sig. (2-tailed)	.132	.000	.000	.000	.000	.000	.000	
	N	2432	2487	2487	2487	2487	2487	2487	2487

**. Correlation is significant at the 0.01 level (2-tailed).

Away Team Wins and Away Team Stadium Orientation

Correlations

		Away_Team_Win	AO_NWSE	AO_SN	AO_SWNE	AO_WE
Away_Team_Win	Pearson Correlation	1	-.054**	.031	.037	-.025
	Sig. (2-tailed)		.007	.122	.070	.210
	N	2434	2434	2434	2434	2434
AO_NWSE	Pearson Correlation	-.054**	1	-.273**	-.409**	-.225**
	Sig. (2-tailed)	.007		.000	.000	.000
	N	2434	2487	2487	2487	2487
AO_SN	Pearson Correlation	.031	-.273**	1	-.448**	-.247**
	Sig. (2-tailed)	.122	.000		.000	.000
	N	2434	2487	2487	2487	2487
AO_SWNE	Pearson Correlation	.037	-.409**	-.448**	1	-.369**
	Sig. (2-tailed)	.070	.000	.000		.000
	N	2434	2487	2487	2487	2487
AO_WE	Pearson Correlation	-.025	-.225**	-.247**	-.369**	1
	Sig. (2-tailed)	.210	.000	.000	.000	
	N	2434	2487	2487	2487	2487

**. Correlation is significant at the 0.01 level (2-tailed).

Home Team Wins and Away Team Stadium Orientation

Correlations

		Home_ Team_Win	AO_NWSE	AO_SN	AO_SWNE	AO_WE
Home_ Team_Win	Pearson Correlation	1	.054**	-.034	-.050*	.046*
	Sig. (2-tailed)		.008	.095	.014	.023
	N	2432	2432	2432	2432	2432
AO_NWSE	Pearson Correlation	.054**	1	-.273**	-.409**	-.225**
	Sig. (2-tailed)	.008		.000	.000	.000
	N	2432	2487	2487	2487	2487
AO_SN	Pearson Correlation	-.034	-.273**	1	-.448**	-.247**
	Sig. (2-tailed)	.095	.000		.000	.000
	N	2432	2487	2487	2487	2487
AO_SWNE	Pearson Correlation	-.050*	-.409**	-.448**	1	-.369**
	Sig. (2-tailed)	.014	.000	.000		.000
	N	2432	2487	2487	2487	2487
AO_WE	Pearson Correlation	.046*	-.225**	-.247**	-.369**	1
	Sig. (2-tailed)	.023	.000	.000	.000	
	N	2432	2487	2487	2487	2487

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Away Team Wins and Away Team Division

Correlations

		Away_Team_Win	A_Divison_C	A_Divison_E	A_Divison_W
Away_Team_Win	Pearson Correlation	1	-.048*	-.005	.054**
	Sig. (2-tailed)		.017	.794	.008
	N	2434	2434	2434	2434
A_Divison_C	Pearson Correlation	-.048*	1	-.505**	-.496**
	Sig. (2-tailed)	.017		.000	.000
	N	2434	2487	2487	2487
A_Divison_E	Pearson Correlation	-.005	-.505**	1	-.500**
	Sig. (2-tailed)	.794	.000		.000
	N	2434	2487	2487	2487
A_Divison_W	Pearson Correlation	.054**	-.496**	-.500**	1
	Sig. (2-tailed)	.008	.000	.000	
	N	2434	2487	2487	2487

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Home Team Wins and Away Team Division

Correlations

		Home_ Team_Win	A_Divison_C	A_Divison_E	A_Divison_W
Home_ Team_Win	Pearson Correlation	1	.045*	.022	-.067**
	Sig. (2-tailed)		.027	.275	.001
	N	2432	2432	2432	2432
A_Divison_C	Pearson Correlation	.045*	1	-.505**	-.496**
	Sig. (2-tailed)	.027		.000	.000
	N	2432	2487	2487	2487
A_Divison_E	Pearson Correlation	.022	-.505**	1	-.500**
	Sig. (2-tailed)	.275	.000		.000
	N	2432	2487	2487	2487
A_Divison_W	Pearson Correlation	-.067**	-.496**	-.500**	1
	Sig. (2-tailed)	.001	.000	.000	
	N	2432	2487	2487	2487

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Correlation between Variables for National Conference

Away Team Wins and Away Team Stadium Orientation

Correlations

		Away_Team_Win	AO_NWSE	AO_SN	AO_SWNE	AO_WE
Away_Team_Win	Pearson Correlation	1	-.050	.003	.109**	-.082**
	Sig. (2-tailed)		.102	.914	.000	.007
	N	1066	1066	1066	1066	1066
AO_NWSE	Pearson Correlation	-.050	1	-.297**	-.355**	-.250**
	Sig. (2-tailed)	.102		.000	.000	.000
	N	1066	1089	1089	1089	1089
AO_SN	Pearson Correlation	.003	-.297**	1	-.426**	-.299**
	Sig. (2-tailed)	.914	.000		.000	.000
	N	1066	1089	1089	1089	1089
AO_SWNE	Pearson Correlation	.109**	-.355**	-.426**	1	-.358**
	Sig. (2-tailed)	.000	.000	.000		.000
	N	1066	1089	1089	1089	1089
AO_WE	Pearson Correlation	-.082**	-.250**	-.299**	-.358**	1
	Sig. (2-tailed)	.007	.000	.000	.000	
	N	1066	1089	1089	1089	1089

**. Correlation is significant at the 0.01 level (2-tailed).

Home Team Wins and Away Team Stadium Orientation

Correlations

		Home_ Team_Win	AO_NWSE	AO_SN	AO_SWNE	AO_WE
Home_ Team_Win	Pearson Correlation	1	.048	.008	-.116**	.081**
	Sig. (2-tailed)		.118	.804	.000	.008
	N	1065	1065	1065	1065	1065
AO_NWSE	Pearson Correlation	.048	1	-.297**	-.355**	-.250**
	Sig. (2-tailed)	.118		.000	.000	.000
	N	1065	1089	1089	1089	1089
AO_SN	Pearson Correlation	.008	-.297**	1	-.426**	-.299**
	Sig. (2-tailed)	.804	.000		.000	.000
	N	1065	1089	1089	1089	1089
AO_SWNE	Pearson Correlation	-.116**	-.355**	-.426**	1	-.358**
	Sig. (2-tailed)	.000	.000	.000		.000
	N	1065	1089	1089	1089	1089
AO_WE	Pearson Correlation	.081**	-.250**	-.299**	-.358**	1
	Sig. (2-tailed)	.008	.000	.000	.000	
	N	1065	1089	1089	1089	1089

**. Correlation is significant at the 0.01 level (2-tailed).

Correlation between Variables for American Conference:

Away Team Wins and Away Team Time Zone Differences

Correlations

		Away_Team_Win	A_TZ_P3	A_TZ_P2	A_TZ_P1	A_TZ_0	A_TZ_N1	A_TZ_N2	A_TZ_N3
Away_Team_Win	Pearson Correlation	1	-.034	-.035	.072*	-.054	-.020	.069*	.030
	Sig. (2-tailed)		.264	.256	.018	.077	.519	.024	.329
	N	1066	1066	1066	1066	1066	1066	1066	1066
A_TZ_P3	Pearson Correlation	-.034	1	-.077*	-.105**	-.223**	-.107**	-.078*	-.068*
	Sig. (2-tailed)	.264		.011	.001	.000	.000	.010	.025
	N	1066	1090	1090	1090	1090	1090	1090	1090
A_TZ_P2	Pearson Correlation	-.035	-.077*	1	-.119**	-.253**	-.121**	-.088**	-.077*
	Sig. (2-tailed)	.256	.011		.000	.000	.000	.004	.011
	N	1066	1090	1090	1090	1090	1090	1090	1090
A_TZ_P1	Pearson Correlation	.072*	-.105**	-.119**	1	-.347**	-.166**	-.121**	-.106**
	Sig. (2-tailed)	.018	.001	.000		.000	.000	.000	.000
	N	1066	1090	1090	1090	1090	1090	1090	1090
A_TZ_0	Pearson Correlation	-.054	-.223**	-.253**	-.347**	1	-.354**	-.258**	-.225**
	Sig. (2-tailed)	.077	.000	.000	.000		.000	.000	.000
	N	1066	1090	1090	1090	1090	1090	1090	1090
A_TZ_N1	Pearson Correlation	-.020	-.107**	-.121**	-.166**	-.354**	1	-.124**	-.108**
	Sig. (2-tailed)	.519	.000	.000	.000	.000		.000	.000
	N	1066	1090	1090	1090	1090	1090	1090	1090
A_TZ_N2	Pearson Correlation	.069*	-.078*	-.088**	-.121**	-.258**	-.124**	1	-.079**
	Sig. (2-tailed)	.024	.010	.004	.000	.000	.000		.009
	N	1066	1090	1090	1090	1090	1090	1090	1090
A_TZ_N3	Pearson Correlation	.030	-.068*	-.077*	-.106**	-.225**	-.108**	-.079**	1
	Sig. (2-tailed)	.329	.025	.011	.000	.000	.000	.009	
	N	1066	1090	1090	1090	1090	1090	1090	1090

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Home Team Wins and Away Team Time Zone Differences

Correlations

		Home_ Team_Win	A_TZ_P3	A_TZ_P2	A_TZ_P1	A_TZ_0	A_TZ_N1	A_TZ_N2	A_TZ_N3
Home_ Team_Win	Pearson Correlation	1	.036	.036	-.038	.048	.015	-.088**	-.041
	Sig. (2-tailed)		.246	.247	.215	.119	.616	.004	.183
	N	1065	1065	1065	1065	1065	1065	1065	1065
A_TZ_P3	Pearson Correlation	.036	1	-.077*	-.105**	-.223**	-.107**	-.078*	-.068*
	Sig. (2-tailed)	.246		.011	.001	.000	.000	.010	.025
	N	1065	1090	1090	1090	1090	1090	1090	1090
A_TZ_P2	Pearson Correlation	.036	-.077*	1	-.119**	-.253**	-.121**	-.088**	-.077*
	Sig. (2-tailed)	.247	.011		.000	.000	.000	.004	.011
	N	1065	1090	1090	1090	1090	1090	1090	1090
A_TZ_P1	Pearson Correlation	-.038	-.105**	-.119**	1	-.347**	-.166**	-.121**	-.106**
	Sig. (2-tailed)	.215	.001	.000		.000	.000	.000	.000
	N	1065	1090	1090	1090	1090	1090	1090	1090
A_TZ_0	Pearson Correlation	.048	-.223**	-.253**	-.347**	1	-.354**	-.258**	-.225**
	Sig. (2-tailed)	.119	.000	.000	.000		.000	.000	.000
	N	1065	1090	1090	1090	1090	1090	1090	1090
A_TZ_N1	Pearson Correlation	.015	-.107**	-.121**	-.166**	-.354**	1	-.124**	-.108**
	Sig. (2-tailed)	.616	.000	.000	.000	.000		.000	.000
	N	1065	1090	1090	1090	1090	1090	1090	1090
A_TZ_N2	Pearson Correlation	-.088**	-.078*	-.088**	-.121**	-.258**	-.124**	1	-.079**
	Sig. (2-tailed)	.004	.010	.004	.000	.000	.000		.009
	N	1065	1090	1090	1090	1090	1090	1090	1090
A_TZ_N3	Pearson Correlation	-.041	-.068*	-.077*	-.106**	-.225**	-.108**	-.079**	1
	Sig. (2-tailed)	.183	.025	.011	.000	.000	.000	.009	
	N	1065	1090	1090	1090	1090	1090	1090	1090

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Away Team Wins and Home Team Stadium Orientation

Correlations

		Away_Team_ Win	HO_NWSE	HO_SN	HO_SWNE	HO_WE
Away_Team_Win	Pearson Correlation	1	.075*	.034	-.048	-.059
	Sig. (2-tailed)		.014	.268	.119	.056
	N	1066	1066	1066	1066	1066
HO_NWSE	Pearson Correlation	.075*	1	-.250**	-.467**	-.199**
	Sig. (2-tailed)	.014		.000	.000	.000
	N	1066	1090	1090	1090	1090
HO_SN	Pearson Correlation	.034	-.250**	1	-.464**	-.198**
	Sig. (2-tailed)	.268	.000		.000	.000
	N	1066	1090	1090	1090	1090
HO_SWNE	Pearson Correlation	-.048	-.467**	-.464**	1	-.369**
	Sig. (2-tailed)	.119	.000	.000		.000
	N	1066	1090	1090	1090	1090
HO_WE	Pearson Correlation	-.059	-.199**	-.198**	-.369**	1
	Sig. (2-tailed)	.056	.000	.000	.000	
	N	1066	1090	1090	1090	1090

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Away Team Win and Away Team Division

Correlations

		Away_Team_Win	A_Divison_C	A_Divison_E	A_Divison_W
Away_Team_Win	Pearson Correlation	1	-.090**	-.005	.095**
	Sig. (2-tailed)		.003	.869	.002
	N	1066	1066	1066	1066
A_Divison_C	Pearson Correlation	-.090**	1	-.503**	-.496**
	Sig. (2-tailed)	.003		.000	.000
	N	1066	1090	1090	1090
A_Divison_E	Pearson Correlation	-.005	-.503**	1	-.500**
	Sig. (2-tailed)	.869	.000		.000
	N	1066	1090	1090	1090
A_Divison_W	Pearson Correlation	.095**	-.496**	-.500**	1
	Sig. (2-tailed)	.002	.000	.000	
	N	1066	1090	1090	1090

**. Correlation is significant at the 0.01 level (2-tailed).

Away Team Wins and Away Team Roof Type

Correlations

		Away_Team_Win	A_Roof_Type_O	A_Roof_Type_F	A_Roof_Type_R
Away_Team_Win	Pearson Correlation	1	-.071*	.011	.071*
	Sig. (2-tailed)		.021	.726	.020
	N	1066	1066	1066	1066
A_Roof_Type_O	Pearson Correlation	-.071*	1	-.444**	-.829**
	Sig. (2-tailed)	.021		.000	.000
	N	1066	1090	1090	1090
A_Roof_Type_F	Pearson Correlation	.011	-.444**	1	-.134**
	Sig. (2-tailed)	.726	.000		.000
	N	1066	1090	1090	1090
A_Roof_Type_R	Pearson Correlation	.071*	-.829**	-.134**	1
	Sig. (2-tailed)	.020	.000	.000	
	N	1066	1090	1090	1090

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Home Team Wins and Home Team Stadium Orientation

Correlations

		Home_ Team_Win	HO_NWSE	HO_SN	HO_SWNE	HO_WE
Home_ Team_Win	Pearson Correlation	1	-.100**	-.001	.038	.062*
	Sig. (2-tailed)		.001	.976	.210	.043
	N	1065	1065	1065	1065	1065
HO_NWSE	Pearson Correlation	-.100**	1	-.250**	-.467**	-.199**
	Sig. (2-tailed)	.001		.000	.000	.000
	N	1065	1090	1090	1090	1090
HO_SN	Pearson Correlation	-.001	-.250**	1	-.464**	-.198**
	Sig. (2-tailed)	.976	.000		.000	.000
	N	1065	1090	1090	1090	1090
HO_SWNE	Pearson Correlation	.038	-.467**	-.464**	1	-.369**
	Sig. (2-tailed)	.210	.000	.000		.000
	N	1065	1090	1090	1090	1090
HO_WE	Pearson Correlation	.062*	-.199**	-.198**	-.369**	1
	Sig. (2-tailed)	.043	.000	.000	.000	
	N	1065	1090	1090	1090	1090

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Home Team Wins and Away Team Roof Type

Correlations

		Home_ Team_Win	A_Roof_Type _O	A_Roof_Type _F	A_Roof_Type _R
Home_ Team_Win	Pearson Correlation	1	.096**	-.018	-.095**
	Sig. (2-tailed)		.002	.555	.002
	N	1065	1065	1065	1065
A_Roof_Type_O	Pearson Correlation	.096**	1	-.444**	-.829**
	Sig. (2-tailed)	.002		.000	.000
	N	1065	1090	1090	1090
A_Roof_Type_F	Pearson Correlation	-.018	-.444**	1	-.134**
	Sig. (2-tailed)	.555	.000		.000
	N	1065	1090	1090	1090
A_Roof_Type_R	Pearson Correlation	-.095**	-.829**	-.134**	1
	Sig. (2-tailed)	.002	.000	.000	
	N	1065	1090	1090	1090

**. Correlation is significant at the 0.01 level (2-tailed).

Home Team Wins and Away Team Stadium Orientation

Correlations

		Home_ Team_Win	AO_NWSE	AO_SN	AO_SWNE	AO_WE
Home_ Team_Win	Pearson Correlation	1	.051	-.076*	.014	.010
	Sig. (2-tailed)		.098	.013	.652	.755
	N	1065	1065	1065	1065	1065
AO_NWSE	Pearson Correlation	.051	1	-.248**	-.468**	-.197**
	Sig. (2-tailed)	.098		.000	.000	.000
	N	1065	1090	1090	1090	1090
AO_SN	Pearson Correlation	-.076*	-.248**	1	-.466**	-.196**
	Sig. (2-tailed)	.013	.000		.000	.000
	N	1065	1090	1090	1090	1090
AO_SWNE	Pearson Correlation	.014	-.468**	-.466**	1	-.370**
	Sig. (2-tailed)	.652	.000	.000		.000
	N	1065	1090	1090	1090	1090
AO_WE	Pearson Correlation	.010	-.197**	-.196**	-.370**	1
	Sig. (2-tailed)	.755	.000	.000	.000	
	N	1065	1090	1090	1090	1090

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Appendix C: LR Tests Results

1. One Single Season LR Analysis

1.1. Single Season of NFL (2016)

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	25.997	23	.301
	Block	25.997	23	.301
	Model	25.997	23	.301

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	322.618[a]	.097	.130

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	9.932	8	.270

Contingency Table for Hosmer and Lemeshow Test

		Home_Team_Win = 0		Home_Team_Win = 1		
		Observed	Expected	Observed	Expected	Total
Step 1	1	21	18.972	5	7.028	26
	2	10	13.153	12	8.847	22
	3	9	13.473	17	12.527	26
	4	13	12.411	14	14.589	27
	5	15	11.095	11	14.905	26
	6	10	9.678	15	15.322	25
	7	10	9.034	16	16.966	26
	8	10	8.011	16	17.989	26
	9	7	7.423	20	19.577	27
	10	3	4.750	22	20.250	25

Classification Table[a]

Observed			Predicted: Home_Team_Win 0	Predicted: Home_Team_Win 1	Percentage Correct
Step 1	Home_Team_Win	0	39	69	36.1
		1	29	119	80.4
	Overall Percentage				61.7

a. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	Travel_WE	-1.873	1.697	1.218	1	.270	.154
	Travel_EW	-.250	.655	.146	1	.702	.779
	ATZD_P3	1.675	1.348	1.545	1	.214	5.341
	ATZD_P2	.919	1.175	.612	1	.434	2.507
	ATZD_N1	1.516	1.531	.980	1	.322	4.552
	ATZD_N2	.472	.973	.235	1	.628	1.603
	H_PST	-1.741	1.608	1.171	1	.279	.175
	H_MST	-1.196	1.443	.688	1	.407	.302
	H_CST	-.037	.630	.004	1	.953	.963
	HO_EW	-1.351	.703	3.692	1	.055	.259
	HO_NESW	-.864	.492	3.092	1	.079	.421
	HO_NS	.121	.344	.123	1	.725	1.128
	HR_O	-.759	.601	1.595	1	.207	.468
	HR_F	-1.297	.661	3.852	1	.050	.273
	HS_G	-.397	.372	1.138	1	.286	.672
	A_PST	1.617	1.556	1.079	1	.299	5.036
	A_MST	.984	1.129	.760	1	.383	2.674
	AO_EW	-.354	.651	.295	1	.587	.702
	AO_NESW	.663	.513	1.671	1	.196	1.940
	AO_NS	.523	.344	2.320	1	.128	1.688
	AR_O	.016	.567	.001	1	.978	1.016
	AR_F	.758	.641	1.398	1	.237	2.134
	AS_G	.578	.366	2.492	1	.114	1.783
	Constant	.998	.799	1.559	1	.212	2.714

a. Variable(s) entered on step 1: Travel_WE, Travel_EW, ATZD_P3, ATZD_P2, ATZD_N1, ATZD_N2, H_PST, H_MST, H_CST, HO_EW, HO_NESW, HO_NS, HR_O, HR_F, HS_G, A_PST, A_MST, AO_EW, AO_NESW, AO_NS, AR_O, AR_F, AS_G.

1.2. Single Season of MLS (2016)

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	14.425	17	.637
	Block	14.425	17	.637
	Model	14.425	17	.637

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	456.903[a]	.042	.055

a. Estimation terminated at iteration number 3 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	12.995	7	.072

Contingency Table for Hosmer and Lemeshow Test

		Home_Team_Win = 0		Home_Team_Win = 1		
		Observed	Expected	Observed	Expected	Total
Step 1	1	22	23.278	14	12.722	36
	2	20	21.036	15	13.964	35
	3	35	33.478	22	23.522	57
	4	17	18.911	17	15.089	34
	5	21	17.229	13	16.771	34
	6	20	13.788	11	17.212	31
	7	12	14.727	23	20.273	35
	8	8	13.985	27	21.015	35
	9	16	14.567	27	28.433	43

Classification Table[a]

			Predicted		
			Home_Team_Win		Percentage Correct
	Observed		0	1	
Step 1	Home_Team_Win	0	109	62	63.7
		1	77	92	54.4
	Overall Percentage				59.1

a. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	WE	2.166	1.285	2.841	1	.092	8.724
	EW	-.870	.659	1.745	1	.187	.419
	ATZD_P3	-1.119	.890	1.581	1	.209	.327
	ATZD_P2	-1.205	.763	2.495	1	.114	.300
	ATZD_N1	-1.535	1.172	1.714	1	.190	.216
	ATZD_N2	-.709	.790	.805	1	.370	.492
	H_PST	2.231	1.246	3.207	1	.073	9.306
	H_MST	1.872	.973	3.701	1	.054	6.503
	H_CST	.842	.597	1.987	1	.159	2.320
	HO_NESW	-.086	.294	.086	1	.769	.917
	HR_O	1.065	.622	2.931	1	.087	2.900
	HS_G	-.579	.364	2.524	1	.112	.561
	A_PST	-2.368	1.156	4.193	1	.041	.094
	A_MST	-1.143	.792	2.084	1	.149	.319
	AO_NESW	.095	.296	.103	1	.749	1.099
	AR_O	.088	.608	.021	1	.885	1.092
	AS_G	-.745	.363	4.201	1	.040	.475
	Constant	-.181	.911	.040	1	.842	.834

a. Variable(s) entered on step 1: WE, EW, ATZD_P3, ATZD_P2, ATZD_N1, ATZD_N2, H_PST, H_MST, H_CST, HO_NESW, HR_O, HS_G, A_PST, A_MST, AO_NESW, AR_O, AS_G.

1.3. Single Season of MLB (2016)

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	45.620	25	.007
	Block	45.620	25	.007
	Model	45.620	25	.007

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	3324.833[a]	.019	.025

a. Estimation terminated at iteration number 3 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	5.897	8	.659

Contingency Table for Hosmer and Lemeshow Test

		Home_ Team_Win = 0		Home_ Team_Win = 1		
		Observed	Expected	Observed	Expected	Total
Step 1	1	152	147.915	92	96.085	244
	2	135	132.972	107	109.028	242
	3	132	133.847	123	121.153	255
	4	124	124.064	120	119.936	244
	5	115	120.557	131	125.443	246
	6	109	116.467	141	133.533	250
	7	105	110.791	144	138.209	249
	8	118	104.674	127	140.326	245
	9	97	100.200	152	148.800	249
	10	82	77.513	128	132.487	210

Classification Table[a]

Observed			Predicted Home_ Team_Win 0	Predicted Home_ Team_Win 1	Percentage Correct
Step 1	Home_ Team_Win	0	539	630	46.1
		1	436	829	65.5
	Overall Percentage				56.2

a. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	Travel_WE	.280	.165	2.877	1	.090	1.323
	Travel_EW	.199	.170	1.364	1	.243	1.220
	A_TZ_P3	1.529	1.191	1.649	1	.199	4.614
	A_TZ_P2	1.420	1.172	1.469	1	.226	4.137
	A_TZ_P1	.991	.769	1.660	1	.198	2.693
	A_TZ_0	1.057	.617	2.938	1	.086	2.877
	A_TZ_N1	.894	.553	2.610	1	.106	2.445
	A_TZ_N2	-.092	.283	.106	1	.744	.912
	H_PST	-.659	.560	1.386	1	.239	.517
	H_MST	-.335	.537	.388	1	.533	.716
	H_CST	-.060	.176	.114	1	.736	.942
	H_Roof_Type_O	-.195	.122	2.561	1	.110	.823
	H_Roof_Type_F	.475	.346	1.882	1	.170	1.608
	H_Playing_Surface_G	.455	.274	2.754	1	.097	1.577
	HO_NWSE	-.104	.142	.534	1	.465	.901
	HO_SN	-.034	.147	.053	1	.818	.967
	HO_SWNE	.058	.125	.212	1	.645	1.059
	A_PDT	.742	.541	1.878	1	.171	2.100
	A_MDT	.263	.478	.302	1	.583	1.300
	A_Roof_Type_O	.277	.122	5.160	1	.023	1.320
	A_Roof_Type_F	-.207	.345	.361	1	.548	.813
	A_Playing_Surface_G	-.472	.273	2.986	1	.084	.624
	AO_NWSE	.066	.143	.217	1	.642	1.069
	AO_SN	-.291	.147	3.909	1	.048	.748
	AO_SWNE	-.314	.126	6.248	1	.012	.731
	Constant	-.900	.731	1.516	1	.218	.407

a. Variable(s) entered on step 1: Travel_WE, Travel_EW, A_TZ_P3, A_TZ_P2, A_TZ_P1, A_TZ_0, A_TZ_N1, A_TZ_N2, H_PST, H_MST, H_CST, H_Roof_Type_O, H_Roof_Type_F, H_Playing_Surface_G, HO_NWSE, HO_SN, HO_SWNE, A_PDT, A_MDT, A_Roof_Type_O, A_Roof_Type_F, A_Playing_Surface_G, AO_NWSE, AO_SN, AO_SWNE.

2. One team LR analysis

2.1. One Team Best Record of the Competition

2.1.1. One Team Best Record of NFL (2018, Rams)

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	11.886	3	.008
	Block	11.886	3	.008
	Model	11.886	3	.008

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	10.044[a]	.524	.703

a. Estimation terminated at iteration number 20 because maximum iterations has been reached. Final solution cannot be found.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	.000	2	1.000

Contingency Table for Hosmer and Lemeshow Test

		Home_Team_Win = 0		Home_Team_Win = 1		
		Observed	Expected	Observed	Expected	Total
Step 1	1	4	4.000	0	.000	4
	2	2	2.000	2	2.000	4
	3	1	1.000	3	3.000	4
	4	0	.000	4	4.000	4

Classification Table[a]

Observed			Predicted Home_Team_Win 0	1	Percentage Correct
Step 1	Home_Team_Win	0	4	3	57.1
		1	0	9	100.0
	Overall Percentage				81.3

a. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	HO_EW	42.406	28420.721	.000	1	.999	2.610E+18
	HO_NS	21.203	20096.485	.000	1	.999	1.615E+9
	AO_NS	-20.104	20096.485	.000	1	.999	.000
	Constant	-21.203	20096.485	.000	1	.999	.000

a. Variable(s) entered on step 1: HO_EW, HO_NS, AO_NS.

2.1.2. One Team Best Record Of MLS (2018, New York Red Bulls)

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	20.400	13	.086
	Block	20.400	13	.086
	Model	20.400	13	.086

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	26.263[a]	.451	.604

a. Estimation terminated at iteration number 20 because maximum iterations has been reached. Final solution cannot be found.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	.000	6	1.000

Contingency Table for Hosmer and Lemeshow Test

		Home_Team_Win = 0		Home_Team_Win = 1		
		Observed	Expected	Observed	Expected	Total
Step 1	1	3	3.000	0	.000	3
	2	2	2.000	0	.000	2
	3	7	7.000	6	6.000	13
	4	2	2.000	2	2.000	4
	5	1	1.000	2	2.000	3
	6	0	.000	4	4.000	4
	7	0	.000	3	3.000	3
	8	0	.000	2	2.000	2

Classification Table[a]

Observed			Predicted Home_Team_Win 0	Predicted Home_Team_Win 1	Percentage Correct
Step 1	Home_Team_Win	0	12	3	80.0
		1	6	13	68.4
	Overall Percentage				73.5

a. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	H_PST(1)	21.128	28406.447	.000	1	.999	1.499E+9
	H_MST(1)	-.154	1.520	.010	1	.919	.857
	H_CST(1)	21.049	40192.962	.000	1	1.000	1.385E+9
	HO_NESW(1)	-21.128	63544.244	.000	1	1.000	.000
	HO_NS(1)	-20.974	63544.244	.000	1	1.000	.000
	HR_O(1)	-21.432	63544.257	.000	1	1.000	.000
	HS_G(1)	21.357	40192.969	.000	1	1.000	1.885E+9
	A_PST(1)	-19.807	18514.239	.000	1	.999	.000
	A_CST(1)	-.154	1.520	.010	1	.919	.857
	AO_NESW(1)	-19.372	47840.389	.000	1	1.000	.000
	AO_NS(1)	1.396	44252.146	.000	1	1.000	4.040
	AR_O(1)	.154	56841.446	.000	1	1.000	1.167
	AS_G(1)	19.807	18514.240	.000	1	.999	399906553
	Constant	18.285	86474.766	.000	1	1.000	87280890.2

a. Variable(s) entered on step 1: H_PST, H_MST, H_CST, HO_NESW, HO_NS, HR_O, HS_G, A_PST, A_CST, AO_NESW, AO_NS, AR_O, AS_G.

2.1.3. One Team Best Record of MLB (2018, Red Sox)

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	27.151	18	.076
	Block	27.151	18	.076
	Model	27.151	18	.076

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	195.425[a]	.154	.207

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	1.090	8	.998

Contingency Table for Hosmer and Lemeshow Test

		Home_ Team_Win = 0		Home_ Team_Win = 1		
		Observed	Expected	Observed	Expected	Total
Step 1	1	16	14.871	4	5.129	20
	2	7	7.398	5	4.602	12
	3	13	12.817	9	9.183	22
	4	7	8.305	8	6.695	15
	5	9	9.541	10	9.459	19
	6	6	6.055	10	9.945	16
	7	5	4.883	10	10.117	15
	8	4	3.759	12	12.241	16
	9	4	3.370	13	13.630	17
	10	1	1.000	9	9.000	10

Classification Table[a]

Observed			Predicted Home_ Team_Win 0	1	Percentage Correct
Step 1	Home_ Team_Win	0	47	25	65.3
		1	31	59	65.6
	Overall Percentage				65.4

a. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	Travel_WE	.561	.817	.471	1	.492	1.752
	Travel_EW	-1.935	1.000	3.740	1	.053	.144
	A_TZ_P3	1.417	1.609	.776	1	.378	4.125
	A_TZ_P1	-.034	1.411	.001	1	.981	.966
	A_TZ_0	-.555	1.100	.255	1	.614	.574
	A_TZ_N1	-.887	1.047	.719	1	.397	.412
	H_Roof_Type_O	.279	.877	.101	1	.751	1.321
	H_Roof_Type_F	2.715	1.427	3.620	1	.057	15.105
	H_Playing_Surface_G	.168	1.223	.019	1	.891	1.183
	HO_NWSE	-.289	1.039	.078	1	.781	.749
	HO_SN	.411	1.026	.160	1	.689	1.508
	HO_SWNE	-1.611	.777	4.297	1	.038	.200
	A_Roof_Type_O	.260	.951	.075	1	.784	1.297
	A_Roof_Type_F	-1.181	1.555	.577	1	.447	.307
	A_Playing_Surface_G	-1.907	1.477	1.667	1	.197	.149
	AO_NWSE	-.888	1.053	.710	1	.399	.412
	AO_SN	-1.257	.961	1.708	1	.191	.285
	AO_SWNE	-1.580	.682	5.364	1	.021	.206
	Constant	4.612	2.452	3.538	1	.060	100.717

a. Variable(s) entered on step 1: Travel_WE, Travel_EW, A_TZ_P3, A_TZ_P1, A_TZ_0, A_TZ_N1, H_Roof_Type_O, H_Roof_Type_F, H_Playing_Surface_G, HO_NWSE, HO_SN, HO_SWNE, A_Roof_Type_O, A_Roof_Type_F, A_Playing_Surface_G, AO_NWSE, AO_SN, AO_SWNE.

2.2. One Team Worst Record

2.2.1. One Team Worst Record of NFL (2018, Cardinals)

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	13.612	13	.402
	Block	13.612	13	.402
	Model	13.612	13	.402

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	8.318[a]	.573	.768

a. Estimation terminated at iteration number 20 because maximum iterations has been reached. Final solution cannot be found.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	.000	6	1.000

Contingency Table for Hosmer and Lemeshow Test

		Home_Team_Win = 0		Home_Team_Win = 1		
		Observed	Expected	Observed	Expected	Total
Step 1	1	2	2.000	0	.000	2
	2	2	2.000	0	.000	2
	3	2	2.000	0	.000	2
	4	1	1.000	1	1.000	2
	5	1	1.000	1	1.000	2
	6	1	1.000	1	1.000	2
	7	0	.000	2	2.000	2
	8	0	.000	2	2.000	2

Classification Table[a]

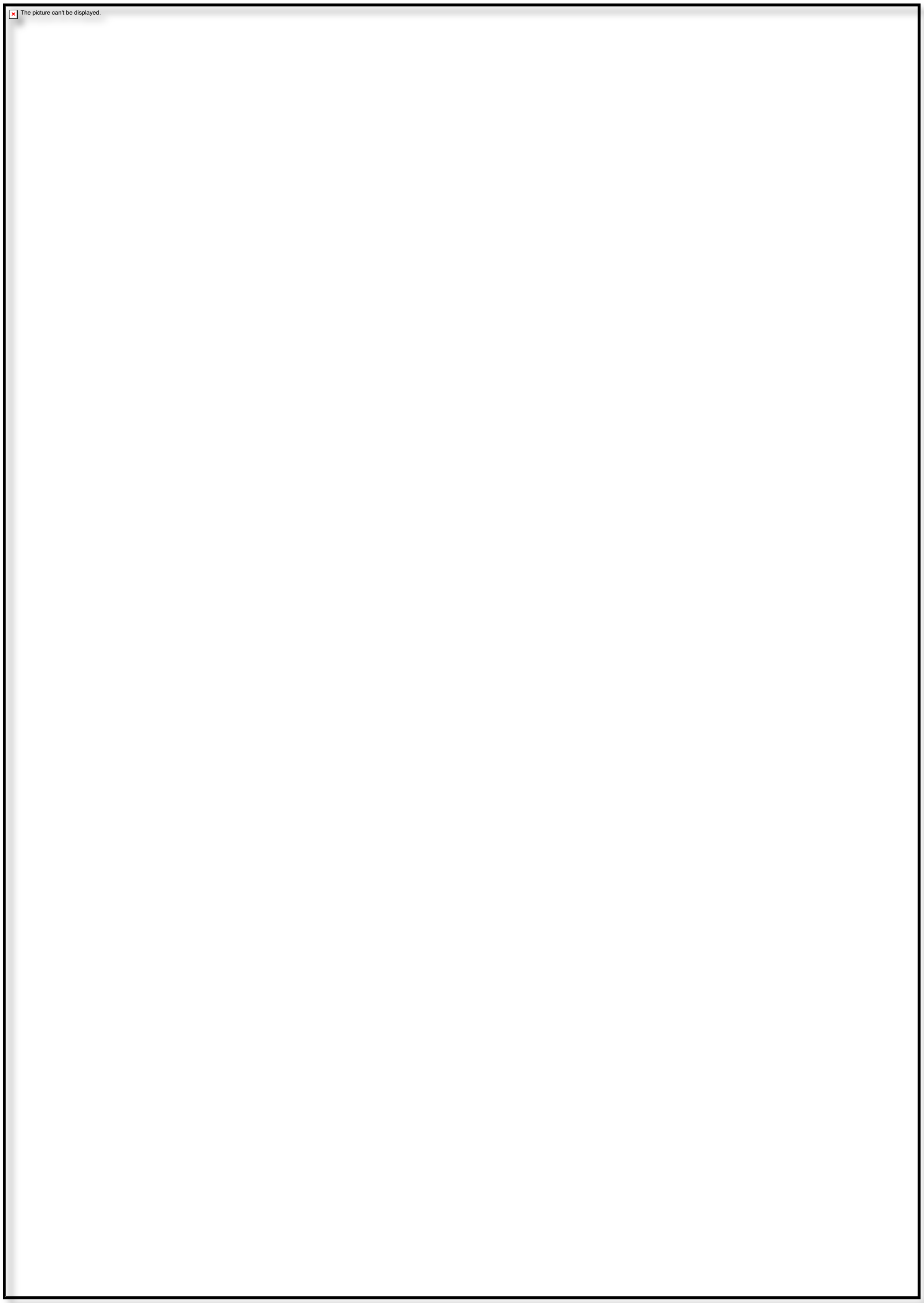
Observed			Predicted Home_Team_Win 0	Predicted Home_Team_Win 1	Percentage Correct
Step 1	Home_Team_Win	0	6	3	66.7
		1	0	7	100.0
	Overall Percentage				81.3

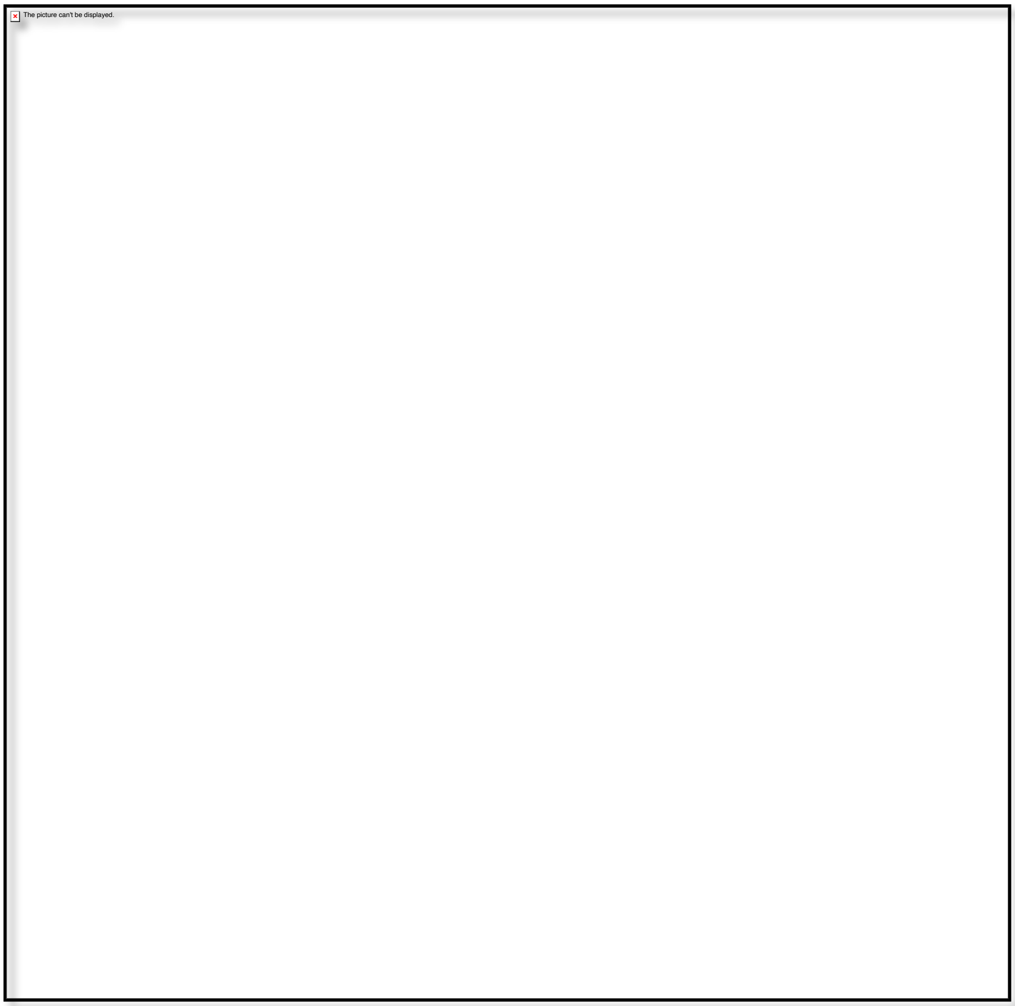
a. The cut value is .500

Variables in the Equation

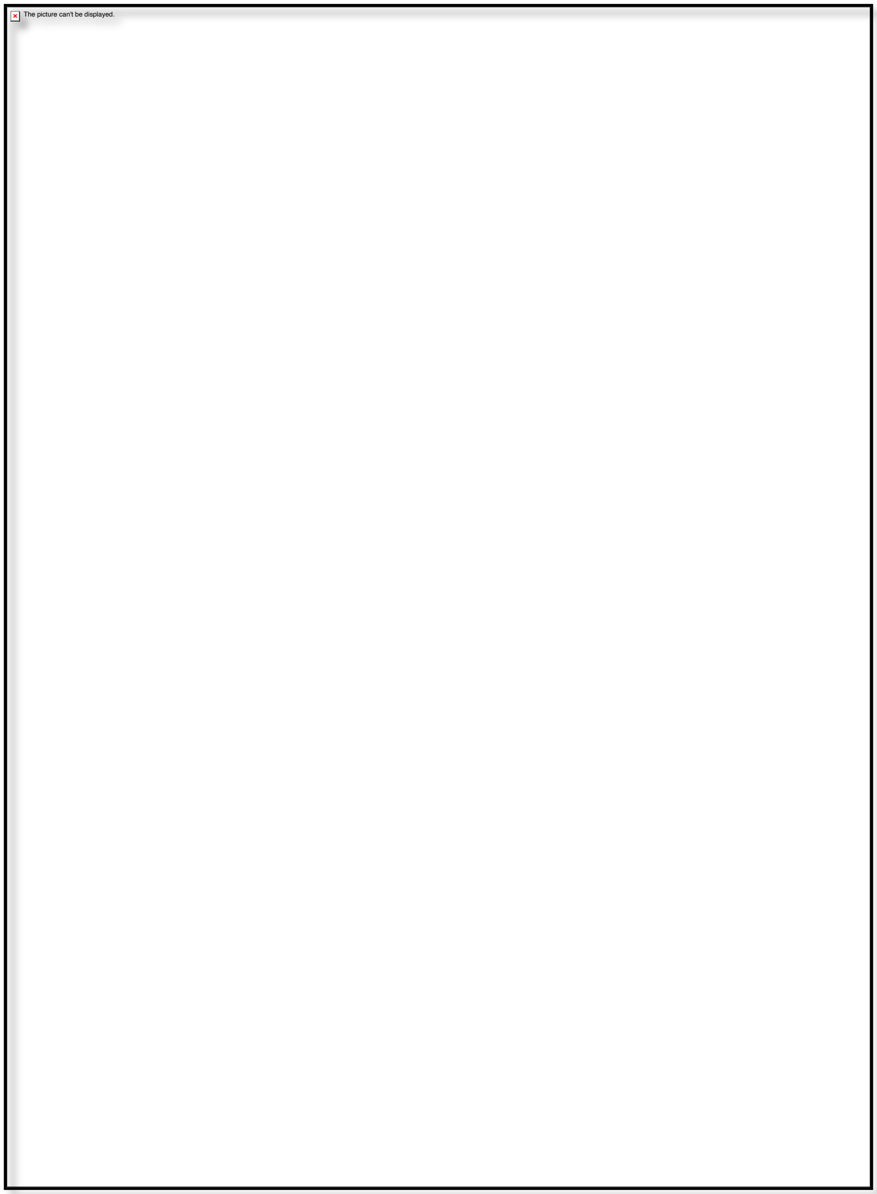
		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	H_PST	21.203	69616.285	.000	1	1.000	1.615E+9

2.2.2. One Team Worst Record of MLS (2018, San Jose)




The picture can't be displayed.

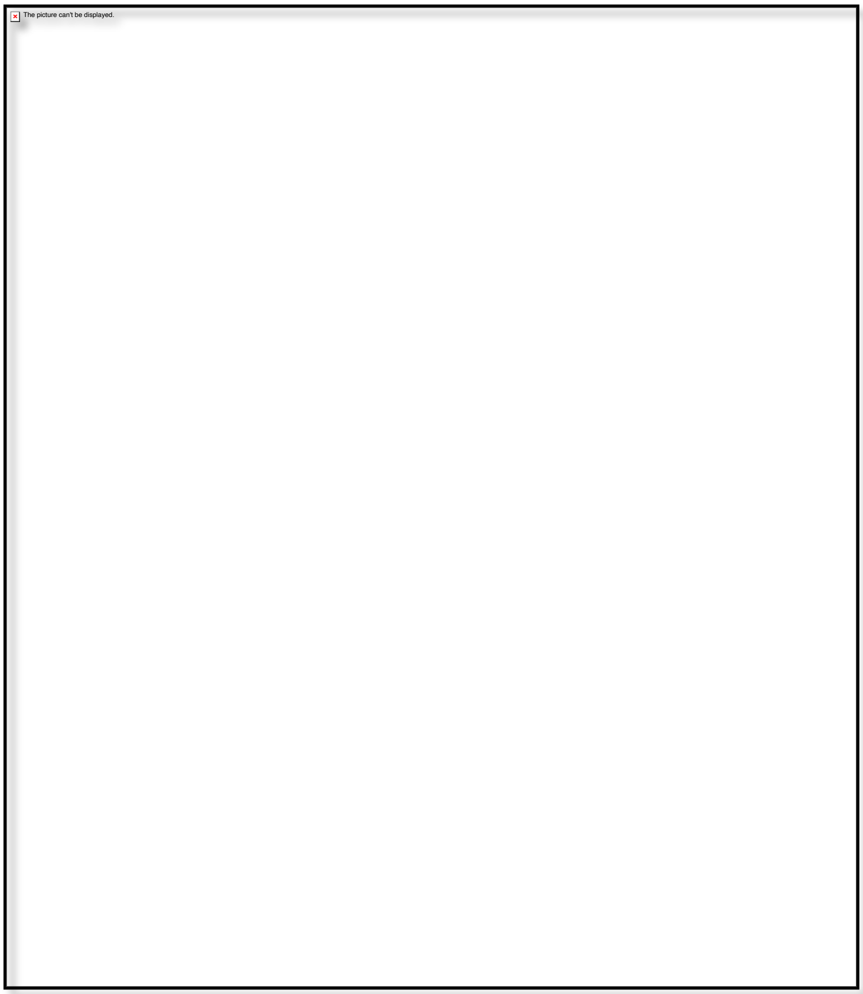
2.2.3. One Team Worst Record of MLB (2018 - Orioles)



The picture can't be displayed.

2.3. One Team Average Record

2.3.1. One Team Average Record of NFL (2018, Packers)



Classification Table[a]

Observed			Predicted Home_Team_Win 0	Predicted Home_Team_Win 1	Percentage Correct
Step 1	Home_Team_Win	0	2	2	50.0
		1	0	12	100.0
	Overall Percentage				87.5

a. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	AR_O	2.398	1.758	1.860	1	.173	11.000
	AR_F	-21.203	28420.722	.000	1	.999	.000
	Constant	.000	1.414	.000	1	1.000	1.000

a. Variable(s) entered on step 1: AR_O, AR_F.

2.3.2. One Team Average Record of MLS (2018, LA Galaxy)

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	18.446	15	.240
	Block	18.446	15	.240
	Model	18.446	15	.240

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	28.217[a]	.419	.561

a. Estimation terminated at iteration number 20 because maximum iterations has been reached. Final solution cannot be found.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	.196	8	1.000

Contingency Table for Hosmer and Lemeshow Test

		Home_Team_Win = 0		Home_Team_Win = 1		
		Observed	Expected	Observed	Expected	Total
Step 1	1	3	3.000	0	.000	3
	2	4	4.000	0	.000	4
	3	2	2.048	1	.952	3
	4	2	1.952	1	1.048	3
	5	1	1.231	1	.769	2
	6	2	1.769	1	1.231	3
	7	1	1.021	1	.979	2
	8	3	3.000	3	3.000	6
	9	1	.979	2	2.021	3
	10	0	.000	5	5.000	5

Classification Table[a]

Observed			Predicted Home_Team_Win 0	1	Percentage Correct
Step 1	Home_Team_Win	0	15	4	78.9
		1	5	10	66.7
	Overall Percentage				73.5

a. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	H_PST(1)	.259	1.280	.041	1	.839	1.296
	H_MST(1)	-21.565	28420.737	.000	1	.999	.000
	H_CST(1)	-.405	1.431	.080	1	.777	.667
	HO_NESW(1)	-21.027	40192.905	.000	1	1.000	.000
	HO_NS(1)	-21.113	40192.905	.000	1	1.000	.000
	HR_O(1)	-20.362	40192.970	.000	1	1.000	.000
	HS_G(1)	-.133	1.445	.008	1	.927	.875
	A_PST(1)	.621	1.448	.184	1	.668	1.861
	A_MST(1)	.000	1.732	.000	1	1.000	1.000
	A_CST(1)	41.000	28858.447	.000	1	.999	6.397E+17
	AO_NESW(1)	-21.455	51976.535	.000	1	1.000	.000
	AO_NS(1)	-1.086	55483.952	.000	1	1.000	.338
	AR_O(1)	-41.000	63747.568	.000	1	.999	.000
	AS_G(1)	42.455	72816.987	.000	1	1.000	2.742E+18
	AS_FT(1)	21.572	77022.731	.000	1	1.000	2.337E+9
	Constant	1.258	138263.207	.000	1	1.000	3.519

a. Variable(s) entered on step 1: H_PST, H_MST, H_CST, HO_NESW, HO_NS, HR_O, HS_G, A_PST, A_MST, A_CST, AO_NESW, AO_NS, AR_O, AS_G, AS_FT.

2.3.3. One Team Average Record of MLB (2018, Diamondbacks)

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	8.382	15	.908
	Block	8.382	15	.908
	Model	8.382	15	.908

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	216.099[a]	.050	.067

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	4.586	8	.801

Contingency Table for Hosmer and Lemeshow Test

		Home_ Team_Win = 0		Home_ Team_Win = 1		
		Observed	Expected	Observed	Expected	Total
Step 1	1	11	11.118	4	3.882	15
	2	12	9.703	3	5.297	15
	3	8	9.087	8	6.913	16
	4	7	6.588	5	5.412	12
	5	4	6.249	8	5.751	12
	6	6	7.273	8	6.727	14
	7	9	8.456	8	8.544	17
	8	7	7.310	9	8.690	16
	9	7	6.913	9	9.087	16
	10	12	10.302	17	18.698	29

Classification Table[a]

Observed			Predicted: Home_ Team_Win 0	1	Percentage Correct
Step 1	Home_ Team_Win	0	48	35	57.8
		1	36	43	54.4
	Overall Percentage				56.2

a. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	Travel_WE	.949	1.291	.540	1	.462	2.582
	A_TZ_P3	2.324	2.456	.895	1	.344	10.215
	A_TZ_P2	1.959	2.348	.696	1	.404	7.094
	A_TZ_P1	2.057	2.419	.723	1	.395	7.819
	A_TZ_0	1.922	1.283	2.245	1	.134	6.833
	A_TZ_N1	1.176	.986	1.421	1	.233	3.240
	A_TZ_N2	1.177	.767	2.357	1	.125	3.244
	H_Roof_Type_O	-.400	.867	.213	1	.645	.670
	HO_NWSE	-.188	.889	.045	1	.832	.828
	HO_SN	-.786	.758	1.076	1	.300	.455
	HO_SWNE	-.662	.739	.804	1	.370	.516
	A_Roof_Type_O	-.555	.703	.623	1	.430	.574
	AO_NWSE	.499	.866	.332	1	.565	1.647
	AO_SN	.346	.738	.220	1	.639	1.413
	AO_SWNE	-.328	.642	.261	1	.609	.720
	Constant	-1.284	2.072	.384	1	.536	.277

a. Variable(s) entered on step 1: Travel_WE, A_TZ_P3, A_TZ_P2, A_TZ_P1, A_TZ_0, A_TZ_N1, A_TZ_N2, H_Roof_Type_O, HO_NWSE, HO_SN, HO_SWNE, A_Roof_Type_O, AO_NWSE, AO_SN, AO_SWNE.

Chapter 4: Direct Application of Research Study Findings to a Practitioner Setting

Abstract

This paper presents an artifact based on findings from Chapter 3. An exploratory data analysis was performed. Various data were gathered to perform the analysis, including match statistics, all team information, travel direction, distance, and all stadium information. Some variables were calculated and added through Esri ArcMap 10.6. The data were initially organized using Microsoft Excel before the data analysis was shifted to IBM's SPSS application. After considering several options, Power BI was the platform selected to create the dashboard. The design science research (DSR) methodology was used for IT artifact creation, review, feedback, and improvement. Based on the data gathered from participants, the dashboard is useable and achieved a positive evaluation.

Introduction

This paper will complete the last two steps of CRISP-DM: evaluation and deployment. This study follows Hevner's (2007) DSR approach. DSR provides a methodology that bridges the technology perspective and behavioral perspective of information systems. DSR is defined as follows:

> Design science research is a research paradigm in which a designer answers questions relevant to human problems via the creation of innovative artifacts, thereby contributing new knowledge to the body of scientific evidence. The designed artifacts are both useful and fundamental in understanding that problem (Hevner & Chatterjee, 2010, p5).

DSR is a suitable and appropriate methodology for IT artifact creation, review, feedback, and improvement. DSR proposes prescriptive artifacts to improve IT performance (Hevner, 2007). IT artifacts can be in the form of constructs and include vocabulary and symbols. DSR models include abstractions or representations, and DSR methods include algorithms or practices. DSR instantiation includes implemented or prototype systems.

The IT artifact proposed in the current study is an instantiation. From the DSR lens, an instantiation is a meta-artifact that can be demonstrated as implemented software or a prototype for problem-solving IT applications. The meta-requirements are the incorporated efforts to transform problems into system objectives (Walls et al., 1992). The proposed artifact is considered as an instantiation artifact, which, in this case, is a web-based application. The artifact has been constructed and assessed. The artifact was built upon the outcome from the LR analysis process.

Method (DSR)

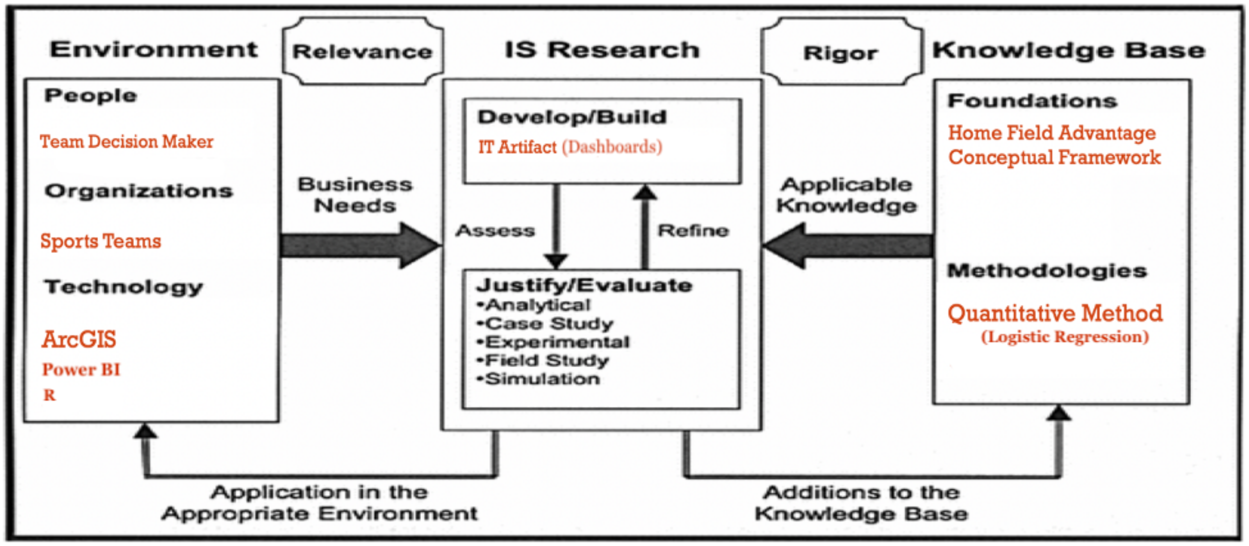
DSR has three iterative and interconnected cycle components: rigor cycle, relevance cycle, and design cycle. The relevance cycle aligns the business requirements with the research objectives. As shown in Figure 4.1, the relevance cycle initiates how DSR applies to the sports industry and team decision-making environment along with research requirements (e.g., problems and opportunities). The relevance cycle begins by determining opportunities, issues, and gaps in an application domain in the environment. Thus, the initial analysis of an environment helps identify gaps; and again, in this case, those regarding spatial, temporal, and stadium attributes.

The rigor cycle focuses on how the artifact and the design process is grounded by scientific foundations, experience, and expertise that informs the study. In this case, the rigor cycle related to the idea of grounding the proposed research topic in scientific knowledge by investigating relevant literature about the competitive intelligence domain and a theoretical foundation such as the HFA conceptual framework. The HFA conceptual framework guided the design and build of the artifact. The rigor cycle supports the design stage in an iterative process. The rigor cycle contributes new knowledge from the present research through rigorous evaluation methods (quantitative, qualitative, or mixed-method approaches) to assess the proposed artifact.

The central design cycle represents the designing, building, and evaluation of the artifact (Hevner & Chatterjee, 2010). Furthermore, the design cycle includes the activities between artifact construction, evaluation, and subsequent feedback to refine the design further. The relevance cycle defined the design as a research criterion for usability, user experience, improvement, and usefulness. The contribution to the design stage is made by providing a web application to show the statistical results of HFA.

The proposed artifact is designed based on these DSR perspectives and built on several iterations for the sake of the refinements and improvement. This artifact will operationalize the proposed model (see Figure 4.1).

Figure 4.1.
DSR Three-Cycle View in the Context of the Study



IT Artifact

The web application (dashboard) was created using Power BI from Microsoft. Power BI is a business analytics service that aims to provide interactive visualizations and business intelligence capabilities with an interface that allows end-users to create their own reports and dashboards. Power BI consists of four elements: a) a Windows desktop application called Power BI Desktop, b) an online SaaS (Software as a Service) service called the Power BI service, c) Power BI mobile apps for Windows, iOS, and Android devices, and d) a Power BI report server. All elements are designed to create, share, and consume business insights effectively (Microsoft, n.d.).

The dashboard provides team decision-makers (coaches, managers, or game analysts) with information to help them understand their opponent's position in the next home or away game. They can then determine which factor will help them avoid losing that game. The output will be displayed in a table of the factors, which explains which factors will have a significant impact on game results. The dashboard also provides statistical information regarding opponents. As shown in Figure 4.2, the main page provides the option to choose which sport to examine. Based on the user choice, the dashboard will direct them to the page where they can test and explore.

To illustrate the capability and features of the dashboard, the user can follow these steps:

- Figure 4.2. shows the main page:

(1) the user has the option to choose one sport to examine by clicking on the sport button.

- Figure 4.3.1, Figure 4.3.2, and Figure 4.3.3., show:

Duration/Time Selection

(1) users can input the temporal duration that they want to examine manually

(2) another option for the user is to choose the duration using a time slider

(3) the user also has the option to choose the season to examine or multiple seasons.

Team Selection

(4) the user has the option to select a team or multiple teams as the home team(s) by clicking on the team button.

(5) the user has the option to select a team or multiple teams as the away team(s) by clicking on the team button.

Features and Results

(6) & (7) cards showing the number of games home and away teams win based on the user selection of previous steps.

(8) shows the logistic regression results (coefficients) based on user selection of the steps from (1) to (5).

(9) charts provide additional information for different features.

Figure 4.2.
Dashboard Main Page

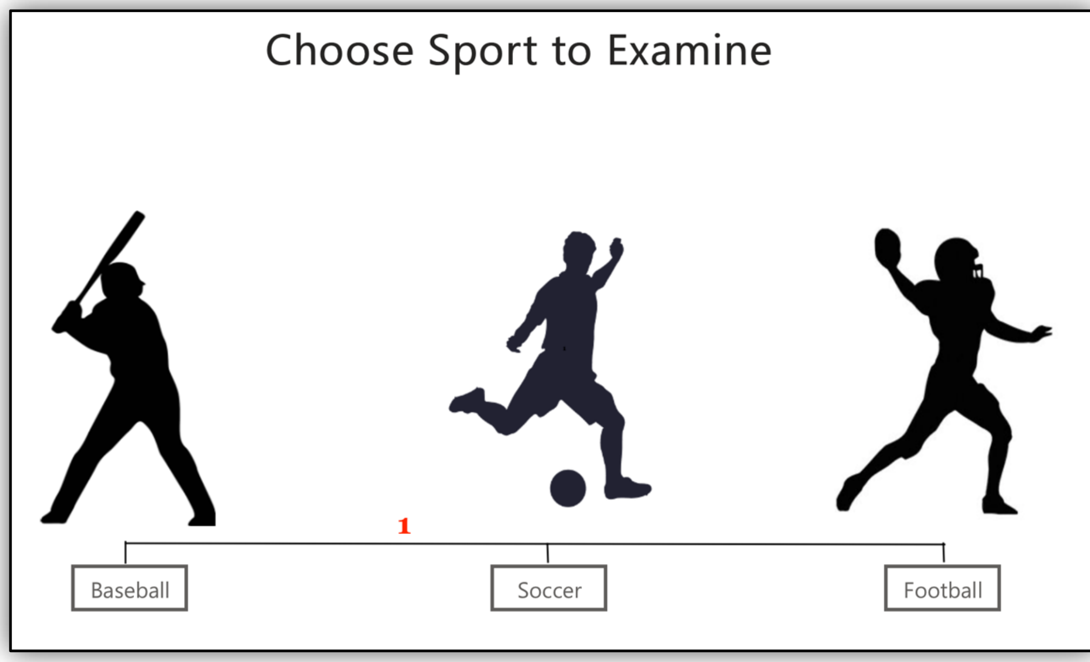
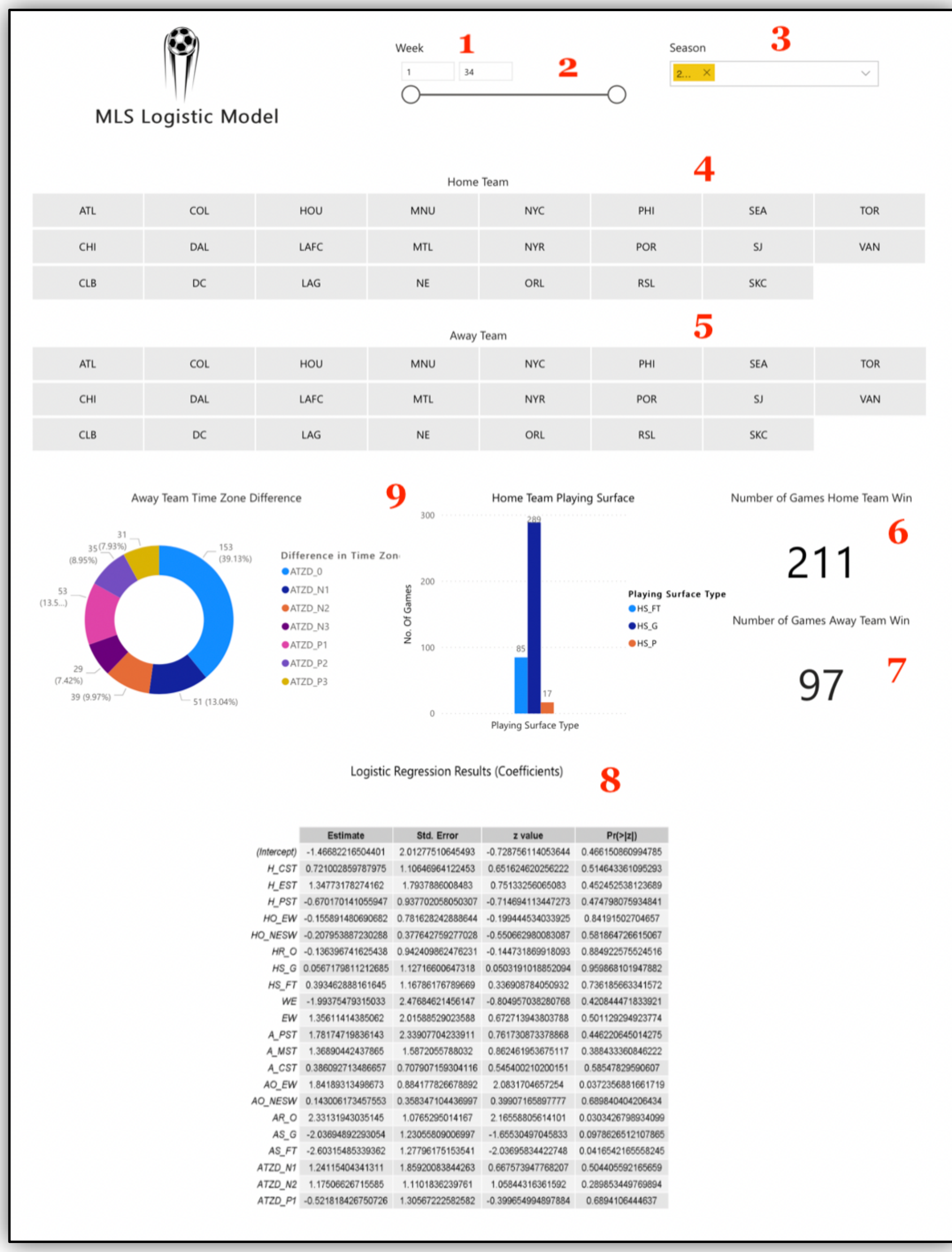


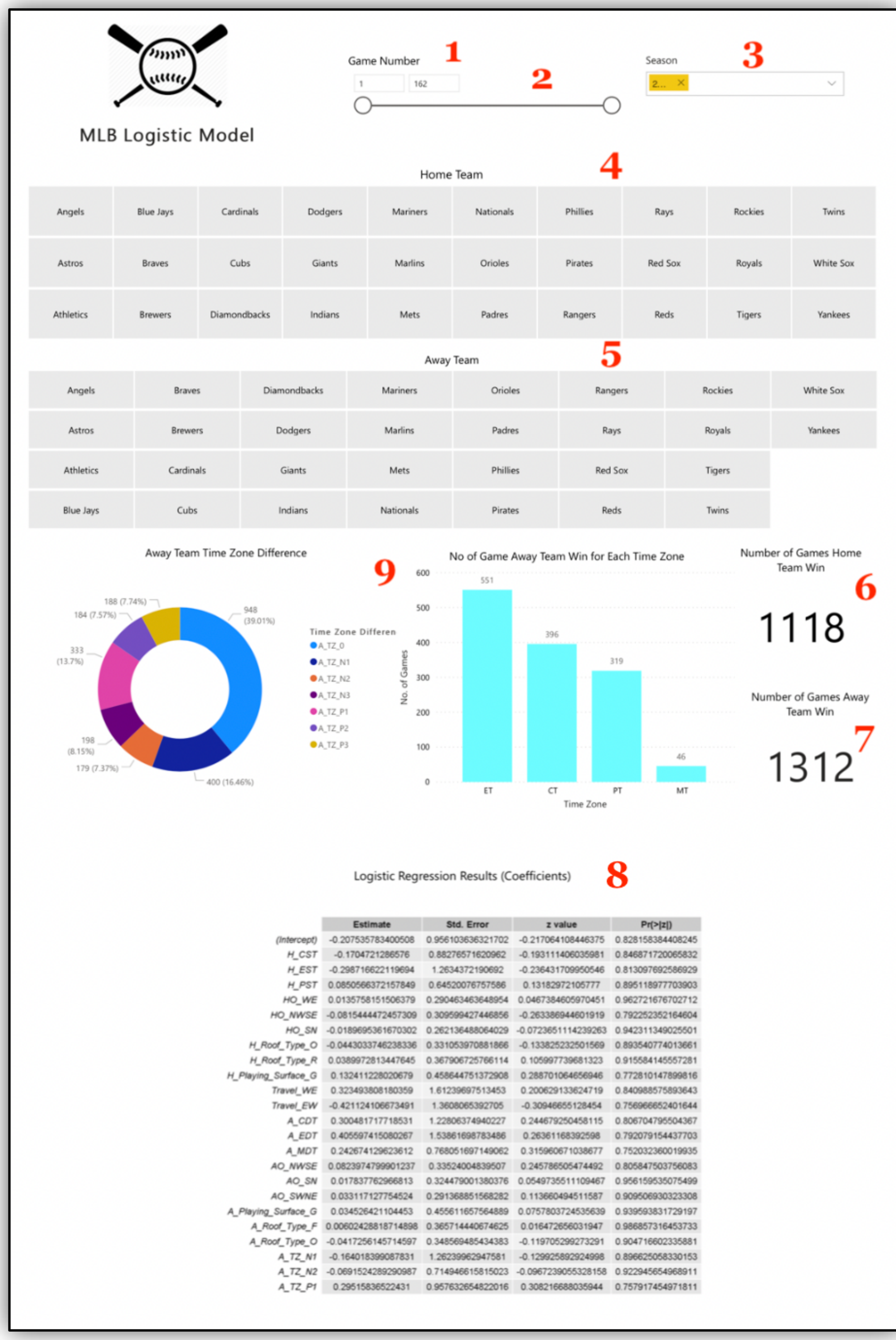
Figure 4.3.1.
Dashboard - MLS Page



	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-1.46682216504401	2.01277510645493	-0.728756114053644	0.466150860994785
H_CST	0.721002859787975	1.10646964122453	0.651624620256222	0.514643361095293
H_EST	1.34773178274162	1.7937886008483	0.75133256065083	0.452452538123689
H_PST	-0.670170141055947	0.937702058050307	-0.714694113447273	0.474798075934841
HO_EW	-0.155891480690682	0.781628242888644	-0.199444534033925	0.84191502704657
HO_NESW	-0.207953887230288	0.377642759277028	-0.550662980083087	0.581864726615067
HR_O	-0.136396741625438	0.942409862476231	-0.144731869918093	0.884922575524516
HS_G	0.0567179811212685	1.12716600647318	0.0503191018852094	0.959868101947882
HS_FT	0.393462888161645	1.16786176789669	0.336908784050932	0.736185663341572
WE	-1.99375479315033	2.47684621456147	-0.804957038280768	0.420844471833921
EW	1.35611414385062	2.01588529023588	0.672713943803788	0.501129294923774
A_PST	1.78174719836143	2.33907704233911	0.761730873378868	0.446220645014275
A_MST	1.36890442437865	1.5872055788032	0.862461953675117	0.388433360846222
A_CST	0.386092713486657	0.707907159304116	0.545400210200151	0.58547829590607
AO_EW	1.84189313498673	0.884177826678892	2.0831704657254	0.0372356881661719
AO_NESW	0.143006173457553	0.358347104436997	0.39907165897777	0.689840404206434
AR_O	2.33131943035145	1.0765295014167	2.16558805614101	0.0303426798934099
AS_G	-2.03694892293054	1.23055809006997	-1.65530497045833	0.0978626512107865
AS_FT	-2.60315485339362	1.27796175153541	-2.03695834422748	0.0416542165558245
ATZD_N1	1.24115404341311	1.85920083844263	0.667573947768207	0.504405592165659
ATZD_N2	1.17506626715585	1.1101836239761	1.05844316361592	0.289853449769894
ATZD_P1	-0.521818426750726	1.30567222582582	-0.399654994897884	0.6894106444637

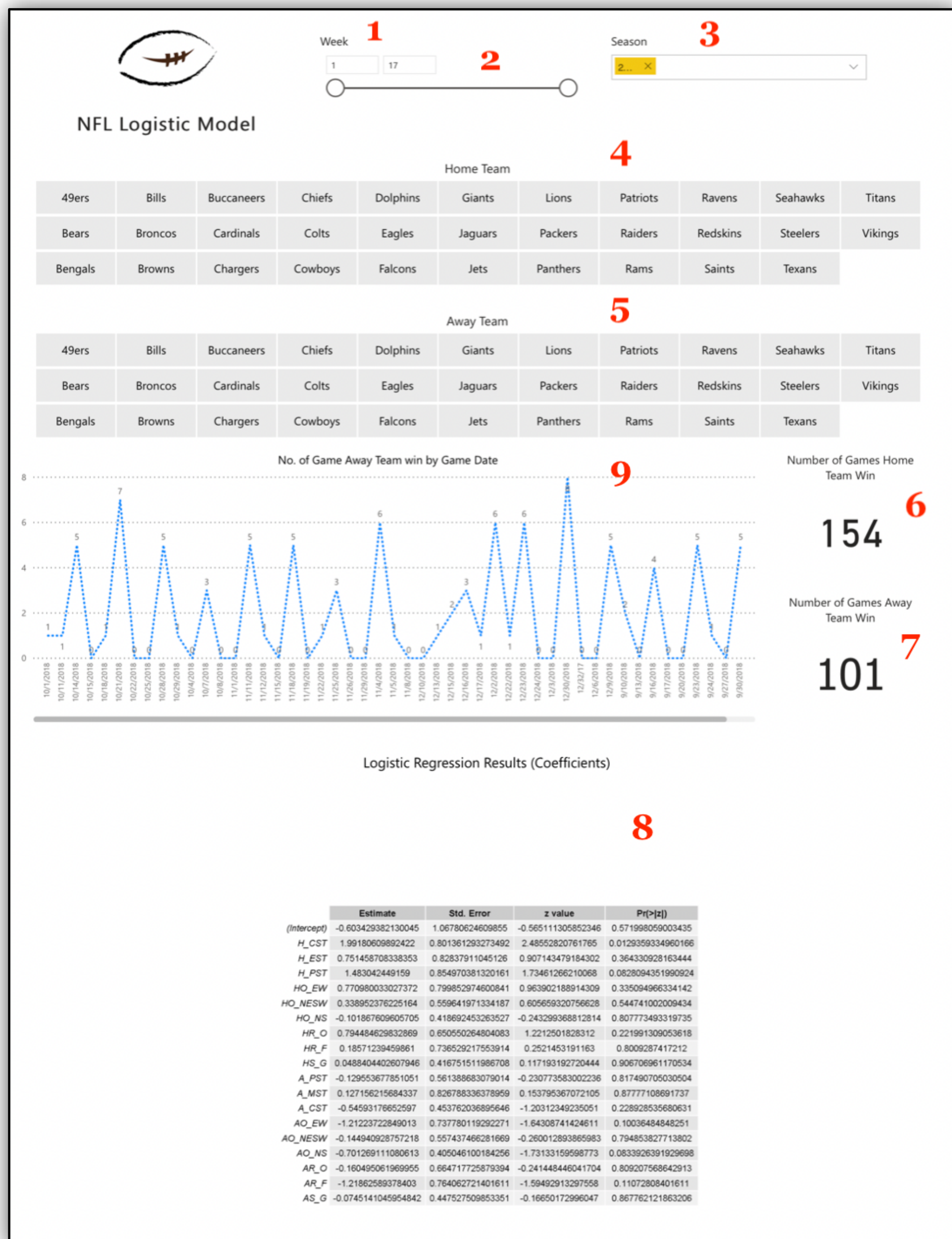
Figure 4.3.2.

Dashboard – MLB Page



	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.207535783400508	0.956103636321702	-0.217064108446375	0.828158384408245
H_CST	-0.1704721286576	0.88276571620962	-0.193111406035981	0.846871720065832
H_EST	-0.298716622119694	1.2634372190692	-0.236431709950546	0.813097692586929
H_PST	0.0850566372157849	0.64520076757586	0.13182972105777	0.895118977703903
HO_WE	0.0135758151506379	0.290463463648954	0.0467384605970451	0.962721676702712
HO_NWSE	-0.0815444472457309	0.309599427446856	-0.263386944601919	0.792252352164604
HO_SN	-0.0189695361670302	0.262136488064029	-0.0723651114239263	0.942311349025501
H_Roof_Type_O	-0.0443033746238336	0.331053970881866	-0.133825232501569	0.893540774013661
H_Roof_Type_R	0.0389972813447645	0.367906725766114	0.105997739681323	0.915584145557281
H_Playing_Surface_G	0.132411228020679	0.458644751372908	0.288701064656946	0.772810147899816
Travel_WE	0.323493808180359	1.61239697513453	0.200629133624719	0.840988575893643
Travel_EW	-0.421124106673491	1.3608065392705	-0.30946655128454	0.756966652401644
A_CDT	0.300481717718531	1.22806374940227	0.244679250458115	0.806704795504367
A_EDT	0.405597415080267	1.53861698783486	0.26361168392598	0.792079154437703
A_MDT	0.242674129623612	0.768051697149062	0.315960671038677	0.752032360019935
AO_NWSE	0.0823974799901237	0.33524004839507	0.245786505474492	0.805847503756083
AO_SN	0.017837762966813	0.324479001380376	0.0549735511109467	0.956159535075499
AO_SWNE	0.033117127754524	0.291368851568282	0.113660494511587	0.909506930323308
A_Playing_Surface_G	0.034526421104453	0.455611657564889	0.0757803724535639	0.939593831729197
A_Roof_Type_F	0.00602428818714898	0.365714440674625	0.016472656031947	0.986857316453733
A_Roof_Type_O	-0.0417256145714597	0.348569485434383	-0.119705299273291	0.904716602335881
A_TZ_N1	-0.164018399087831	1.26239962947581	-0.129925892924998	0.896625058330153
A_TZ_N2	-0.0691524289290987	0.714946615815023	-0.0967239055328158	0.922945654968911
A_TZ_P1	0.29515836522431	0.957632654822016	0.308216688035944	0.757917454971811

Figure 4.3.3.
Dashboard – NFL Page



	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.603429382130045	1.06780624609855	-0.565111305852346	0.571998059003435
H_CST	1.99180609892422	0.801361293273492	2.48552820761765	0.0129359334960166
H_EST	0.751458708338353	0.82837911045126	0.907143479184302	0.364330928163444
H_PST	1.483042449159	0.854970381320161	1.73461266210068	0.0828094351990924
HO_EW	0.770980033027372	0.799852974600841	0.963902188914309	0.335094966334142
HO_NESW	0.338952376225164	0.559641971334187	0.605659320756628	0.544741002009434
HO_NS	-0.101867609605705	0.418692453263527	-0.243299368812814	0.807773493319735
HR_O	0.794484629832869	0.650550264804083	1.2212501828312	0.221991309053618
HR_F	0.18571239459861	0.736529217553914	0.2521453191163	0.8009287417212
HS_G	0.0488404402607946	0.416751511986708	0.117193192720444	0.906706961170534
A_PST	-0.129553677851051	0.561388683079014	-0.230773583002236	0.817490705030504
A_MST	0.127156215684337	0.826788336378959	0.153795367072105	0.87777108691737
A_CST	-0.54593176652597	0.453762036895646	-1.20312349235051	0.228928535680631
AO_EW	-1.21223722849013	0.737780119292271	-1.64308741424611	0.10036484848251
AO_NESW	-0.144940928757218	0.557437466281669	-0.260012893865983	0.794853827713802
AO_NS	-0.701269111080613	0.405046100184256	-1.73133159598773	0.0833926391929698
AR_O	-0.160495061969955	0.664717725879394	-0.241448446041704	0.809207568642913
AR_F	-1.21862589378403	0.764062721401611	-1.59492913297558	0.11072808401611
AS_G	-0.0745141045954842	0.447527509853351	-0.16650172996047	0.867762121863206

Code

As introduced in Chapter 2, the Real Time Learning Machine (RTLM) method was used as a guide for the real-time data mining aspect of this artifact.

Writing and developing the code is done in two stages:

1. Learning Machine (LM) iteration in Python (see Figure 4.4): ML reads from the original table and creates and inserts a new table if the user desires a specific time period.

Figure 4.4.
Machine Learning Iteration in Python

```
import pandas as pd       # read the data from the downloaded CSV file.
data = pd.read_csv('NFL.csv')
# set a numeric id for use as an index for examples.

data=data.head(5)    # I did this to have a smaller size table 500 >5 to show the example
data      # you can see the data here numbers and strings also you can see the indecis we introduce in 1
a=[]                                   # to initialize our output list
for i in range(5):                     # for loop in the size of our table we identify before
    a.append(data.iloc[range(i+1)])         # this tis to add to the empty list in a loop but different one each time
a
a[0]  # this is the output as first element of the list
a[1]  # this is the output as second element of the list, you can see it include 0 and 1 index and so on
a[2]
```

2. LR in R for each sport (see Figure 4.5): The step is important because the NFL, MLS, and MLB have a different number of games for each season. The NFL has 16 games per season, the MLS has 34 games per season, and the MLB has 162 games per season.

Figure 4.5.
LR in R

```
library(caTools)
library(dplyr)
library(rms)


setwd("C:/Users/Abdullah/OneDrive/Documents/R/data")
NFL<- read.csv("NFL_TEST.csv", header = T)

##NFL <- NFL%>% filter(NFL$Week %in% (1:17))


# split the data
split<- sample.split(NFL, SplitRatio = 0.8)
train<- subset(NFL, split== "TRUE")
test<- subset(NFL, split =="FALSE")



mymodel<- glm( Home_Team_Win ~ H_CST + H_EST + H_PST + H_MST + HO_EW + HO_NESW + HO_NS + HO_NWSE +
                 HR_O + HR_F+ HR_R + HS_G + HS_T + A_PST + A_MST + A_CST + A_EST +
                 AO_EW + AO_NESW + AO_NS + AO_NWSE + AR_O + AR_F + AR_R + AS_G + AS_T, data = train,
family='binomial')

summary(mymodel)$coefficients

#run the test data in throuh the model
res <- predict(mymodel, test, type="response")
res

res<- predict(mymodel, train, type="response")
res

#validate the model - confusion matrix
conf<- table(Actual_value=train$Home_Team_Win, Predicted_vale = res >0.5)
conf
#accuracy
accur<- (conf[[1,1]] + conf[[2,2]])/ sum(conf)
accur*100

##R2
R2<- with(summary(mymodel), 1 - deviance/null.deviance)
R2*100
```

Evaluation

To evaluate the dashboard, I used a mixed-method design by conducting a descriptive quantitative analysis through the USE questionnaire adapted from Lund's (2011) USE survey questions, as well as collecting qualitative feedback from dashboard users through open-ended questions attached to the survey. The survey was approved as exempt by the Claremont Graduate University Institutional Review Board (see Appendix D, Appendix E).

The survey was administered using Qualtrics. Thirty student participants were recruited from the Center of Information System and Technology at Claremont Graduate University. The survey was distributed via e-mail to these students (see Appendix F for e-mail sample).

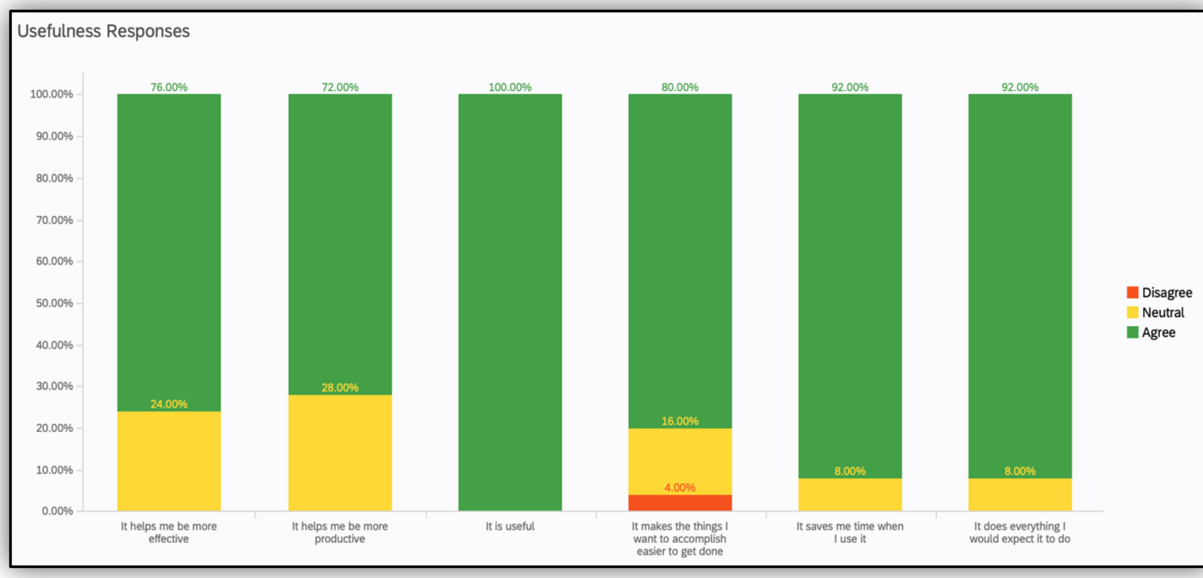
Analysis Results

Based on the data gathered from participants, the dashboard is useable and achieved a positive evaluation. Thirty participants evaluated the dashboard for Usefulness, Ease of Use, Ease of Learning, and User Satisfaction. In addition, the open-ended questions provide insights about user experience and user feedback. Five responses were incomplete; thus, the total sample size of those who completed the survey was 25. The first part of the survey represents the quantitative section of the evaluation.

Usefulness

Most of the respondents agreed that the dashboard helps them to be more effective and productive in making a decision. In addition, they all agreed that the dashboard is useful. Five respondents indicated that the dashboard does not allow them to accomplish what they are looking for easily. The majority of the respondents agreed that the dashboard reduces the time to make a decision while they are using it. Detailed results are shown in Figure 4.6.

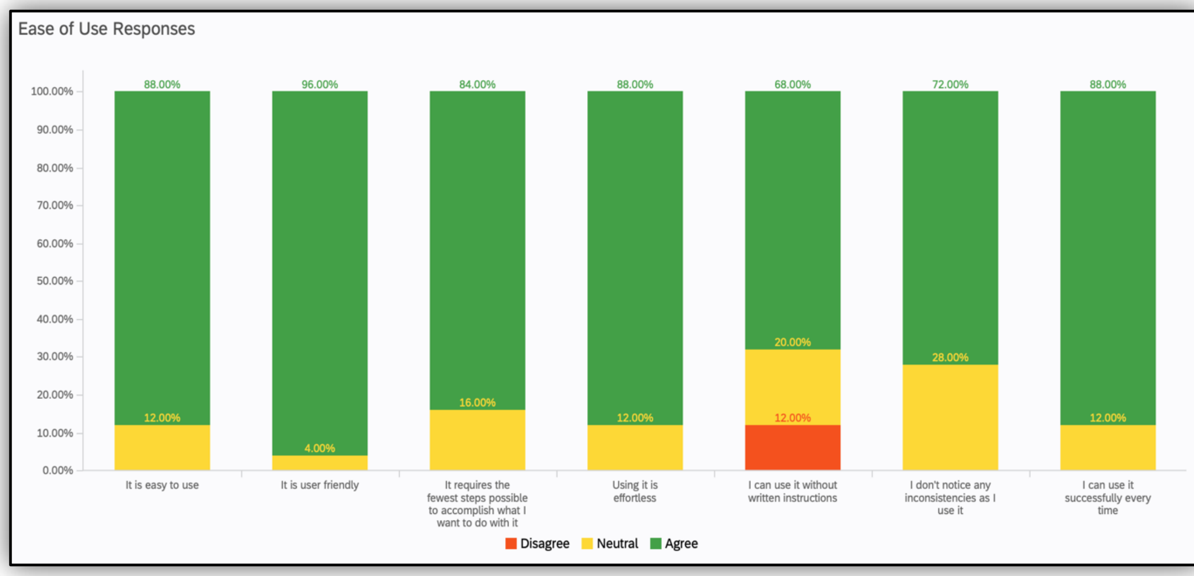
Figure 3.6.
Percentages of responses given to each questionnaire item about the perceived usefulness of the dashboard



Ease of use

The majority of the respondents agreed that the dashboard is easy to use, and it is user friendly. However, 32% of respondents indicated that the dashboard could not be used without written instruction. Future research will include the instructions inside the dashboard or provide a demo option on the main page. Twenty-three respondents agreed that they could use the dashboard successfully every time they use it. Detailed results are shown in Figure 4.7.

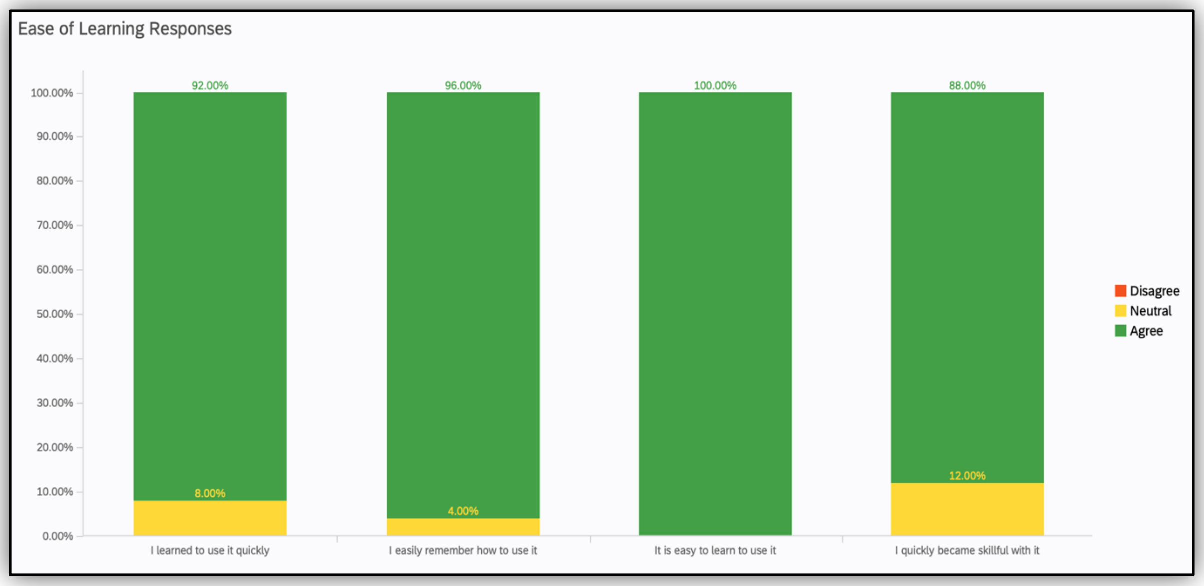
Figure 4.7.
Percentages of responses given to each questionnaire item about the perceived ease of use of the dashboard



Ease of Learning

Based on the responses regarding the ease of learning, the application gains a high rating on all questionnaire items. All respondents agreed that the dashboard is easy to learn and easy to use. Also, twenty-four of the respondents agreed that by using the dashboard they can easily remember how to use it again. Ninety-two percent of the respondents learned how to use the dashboard quickly. Detailed results are shown in Figure 4.8.

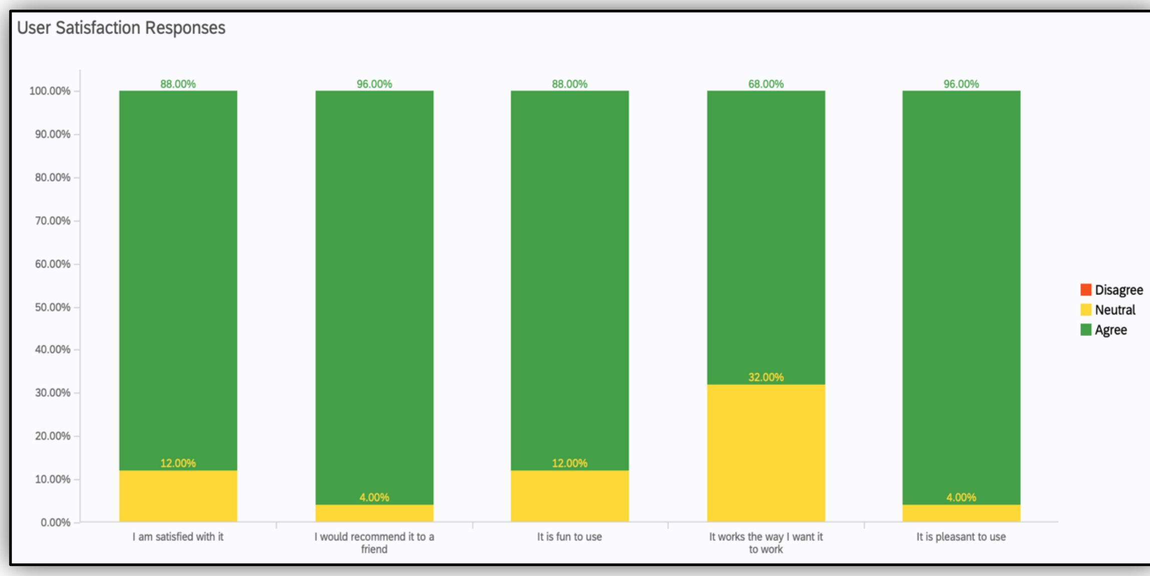
Figure 4.8.
Percentages of responses given to each questionnaire item about the perceived ease of learning of the dashboard



User Satisfaction

Most of the respondents agreed that they were satisfied with the experience they gained by using the application. Ninety-six percent agreed that they would recommend the dashboard to other users. Eight respondents did not agree or disagree that the dashboard worked the way they wanted it to work and provided some feedback. Detailed results are shown in Figure 4.9.

Figure 4.9.
Percentages of responses given to each questionnaire item about the perceived user satisfaction of the dashboard



The second part of the survey represents the qualitative section. To analyze the responses for the open-ended questions, thematic coding was used. Researchers commonly use this analysis approach when analyzing and categorizing qualitative data into summaries of shared meaning patterns called themes (Braun et al., 2018). The analysis approach resulted in two main themes: usefulness and improvements.

Qualitative data offers deep insights that numeric data do not necessarily provide. Participants clearly confirmed the usefulness of the dashboard. They viewed the tool as useful for several tasks in decision-making. For example, one of the participants stated:

> It was helpful to understand how the shown variables explain the winning in sports games.

In a similar manner, another participant stated:

> Listing all variables that effect the home advantage for each game and each team during the season [is] very useful.

Moreover, participants indicated the usability of the dashboard, as quoted below:

> … it was fun to see the home advantage changes during the season
>
> Fun and deliver what I'm looking for
>
> very satisfying and fun
>
> Because the application is organized, it is easy to use it
>
> I think [it] is fun to see the machine learning methods being used in sports.

In contrast, participants suggested ideas and changes to the system that will lead to system improvements. For example, one of the participants suggested:

> Highlight the significance variables that affect HFA.

In a similar manner, another participant suggested:

> Sort the sig. variables from high to low.

In addition, some participants suggested the following:

> Different languages will be useful
>
> add some details for each team and for each League
>
> Highlight the winning and losing numbers with color
>
> Provided with different languages
>
> results table should reduce the numbers after the decimal point

"adding team logos would make it much easier to recognize."

Conclusion

Based on the findings from Chapter 3, this paper explains how an IT artifact was developed and evaluated. The dashboard was created using Power BI from Microsoft. The dashboard provides team decision-makers with information to help them understand their opponent's position in the next home or away game. They can then determine which factor will help them avoid losing that game. The methodology used for IT artifact creation, review, feedback, and improvement is DSR.

Based on the data gathered from participants, the dashboard is useable and achieved a positive evaluation. The quantitative part of the dashboard evaluation assessed the following: usefulness, ease of use, ease of learning, and user satisfaction. In addition, the qualitative part of the evaluation, which are open-ended questions, provided insight regarding user experience, and user feedback. These suggestions and feedback will be included in future iterations of this artifact.

References

Braun, V., Clarke, V., Hayfield, N., & Terry, G. (2018). Thematic analysis. In P. Liamputtong (Ed.), Handbook of research methods in health social sciences (pp. 1–18). https://doi.org/10.1007/978-981-10-2779-6_103-1

Courneya, K. S., & Carron, A. V. (1992). The home advantage in sport competitions: A literature review. *Journal of Sport and Exercise Psychology, 14*(1), 13–27. https://doi./10.4135/9781483332222.n138

Hevner, A. R. (2007). A three cycle view of design science research. *Scandinavian Journal of Information Systems, 19*(2), Article 4. https://www.uio.no/studier/emner/jus/afin/FINF4002/v13/hefner-design.pdf

Hevner, A., & Chatterjee, S. (2010). *Design research in information systems: Theory and practice.* Springer. 10.1007/978-1-4419-5653-8

Lund, A. M. (2001). Measuring usability with the use questionnaire12. Usability interface, 8(2), 3-6. Chicago

Microsoft. (n.d.). *What is power BI?* https://powerbi.microsoft.com/en-us/what-is-power-bi/

Walls, J. G., Widmeyer, G. R., & El Sawy, O. A. (1992). Building an information system design theory for vigilant EIS. *Information Systems Research, 3*(1), 36–59. https://doi.org/10.1287/isre.3.1.36

Appendix D: Institutional Review Board Approval

Claremont Graduate University
Institutional Review Board

02/28/2020

Dear Abdullah,

An IRB representative has conducted a preliminary review of protocol **IRB # 3716 EFFECTIVENESS OF SPATIAL-TEMPORAL DATA USING GIS IN AMERICA'S PROFESSIONAL SPORTS LEAGUES (MLS, MLB, AND NFL).** Pursuant to federal regulations 45 CFR 46.102(e)/(l) your project is not human subjects research, and does not require further IRB review or oversight.

Please note that changes to your protocol may affect this determination. Please contact me directly to discuss any changes you may contemplate.

Respectfully,

James Griffith,
IRB Manager
james.griffith2@cgu.edu

150 East Tenth Street ● Claremont, California 91711-6160
Tel: 909.607.9406

Appendix E: Survey of the Dashboard

Q1: Please rate your experience using the Home Field Advantage in Sports Application with the following perspective:

Usefulness

	Disagree	Neutral	Agree
It helps me be more effective	○	○	○
It helps me be more productive	○	○	○
It is useful	○	○	○
It makes the things I want to accomplish easier to get done	○	○	○
It saves me time when I use it	○	○	○
It does everything I would expect it to do	○	○	○

Q2: Please rate your experience using the Home Field Advantage in Sports Application with the following perspective:

Ease of Use

	Disagree	Neutral	Agree
It is easy to use	○	○	○
It is user friendly	○	○	○
It requires the fewest steps possible to accomplish what I want to do with it	○	○	○
Using it is effortless	○	○	○
I can use it without written instructions	○	○	○
I don't notice any inconsistencies as I use it	○	○	○
I can use it successfully every time	○	○	○

Q3: Please rate your experience using the Home Field Advantage in Sports Application with the following perspective:

Ease of Learning

	Disagree	Neutral	Agree
I learned to use it quickly	○	○	○
I easily remember how to use it	○	○	○
It is easy to learn to use it	○	○	○
I quickly became skillful with it	○	○	○

Q4: Please rate your experience using the Home Field Advantage in Sports Application with the following perspective:

User Satisfaction

	Disagree	Neutral	Agree
I am satisfied with it	○	○	○
I would recommend it to a friend	○	○	○
It is fun to use	○	○	○
It works the way I want it to work	○	○	○
It is pleasant to use	○	○	○

Q5: How does the application provide you with information that helps in making decisions?

Q6: How would you rate your overall experience using this application?

Q7: What can be done to improve your experience?

Q8: Your additional comments and feedback are highly appreciated.

Appendix F: Email sample for evaluation participation request

Dear Participant,

My name is Abdullah Aleissa and I am a doctoral student of information systems and technology at Claremont Graduate University. I would like to invite you to participate in my research study: EFFECTIVENESS OF SPATIAL-TEMPORAL DATA USING GIS IN AMERICA'S PROFESSIONAL SPORTS LEAGUES (MLS, MLB, AND NFL). The purpose of this study is to learn more about the effectiveness and usefulness of spatial-temporal data using GIS in America's professional sports leagues to decision makers by testing different factors including stadium location attributes and result outcomes.

Click on the link below, as you will have access to the application and the whole process will take approximately 10-15 minutes of your time.:

https://app.powerbi.com/Redirect?action=OpenApp&appId=b944b9e9-019f-4275-81e0-74ccfa3e7910&ctid=19afb2c8-5efd-4718-a107-530ed963d11e

After testing the application, please take the survey using the link below:

https://cgu.co1.qualtrics.com/jfe/form/SV_aaWf7Ll4zmppU9f

Please feel free to contact me via email: abdullah.aleissa@cgu.edu if you need any further information or have any questions during this process.

Your participation is highly appreciated. Thank you for your time!

Warmly,

Abdullah Aleissa

Chapter 5: Conclusion

This study incorporated three types of research. First, a literature review was conducted regarding sports analytics and the sports industry in relation to HFA regarding the NFL, MLB, and MLS. Second, a research study was conducted using LR to explore HFA. Lastly, a DSR approach was used to develop an IT artifact to solve real-world problems for professional sports teams.

Discussion

The first paper shared the notable amount of literature on sports analytics and HFA. It was discovered that an insufficient amount of literature exists regarding both sports analytics and HFA, particularly in association with spatial, temporal, and stadium attributes such as field surface type, roof type, time zone, and field orientation for soccer, baseball, and football in the United States.

The second paper discussed how the HFA phenomenon exists in football, baseball, and soccer. An analysis of the spatial-temporal, and stadium attributes found that the LR model was statistically significant and the results vary from one sport to another. Some similarities between variables were significant for all sports; however, differences in variables also existed.

The third paper resulted in a web application (Dashboard) that analyzes data from different sources and enables practitioners to assess the impact of HFA on their team's performance in all three sports.

Contribution

From a design science research perspective, this study seeks to contribute to sports analytics and information technology by studying, understanding, and testing the HFA conceptual framework provided by Courneya and Carron (1992), specifically the game location factor. The study also seeks to provide practical contributions by providing an IT artifact that will help decision-makers in sports teams to win, or avoid losing, on the road.

Contribution to Research (Theoretical Contribution)

This study contributes to the theoretical knowledge base by contributing new knowledge to the information systems field and sports analytics field for scholars who want to develop their knowledge and skills in the DSR methodology. The present study further contributes to the existing literature by investigating the sports analytics domain, and testing and examining how the HFA advantage framework can guide artifact design. Additionally, the study contributes new knowledge through rigorous evaluation methods with respect to system usability, user experience, usefulness, and improvement. While prior literature has shed light on HFA, no prior studies have examined the spatial, temporal, and stadium attributes such as field surface type, roof type, time zone, and field orientation for soccer, baseball, and football in the United States to determine how these elements impact the HFA theoretical framework.

Contribution to Practice (Practical Relevance)

This study presents a designed artifact for practitioners to use to assess the impact of HFA on their team's performance. The artifact was useful in identifying which variables have an impact on game results. The artifact has the potential to help teams improve their winning percentage for specific games.

For scenario purposes, let us assume that the Packers decision-makers (coaches, managers, or game analyst) need to understand their opponents' position in the road game to determine which factors will help the team to avoid losing that game. The decision-makers would see the output as a table of the factors that have a significant impact on game results. In addition, decision-makers will see some statistical information about their opponent.

Another added contribution to practice is the process of designing and building the artifact in an iterative manner. This contribution was inspired by DSR activities, starting with grounding the research ideas drawn from the domain knowledge base to include problems from sports teams in the sports analytics industry. Such a process could spotlight the sequence of studying relevant problems, and conducting analysis, design, and implementation of the artifact for the practitioner to consider as a guideline when doing similar work.

Limitations and Future Research

This research focused on some key aspects of HFA in three major sports in the NFL, MLB, and MLS. Some significant limitations can be addressed in future studies. First, this study focused only on the learning and familiarity factor. However, the result attributes and the data gathered mainly reflect the performance outcomes of teams and do not include data related to the psychological and behavioral states of the players. A player's psychological and behavioral state may impact performance outcomes, and additional studies should consider these factors.

Second, this study only examined each sport from the 2016 season to the 2018 season. The analysis indicates an increase in predictor significance after testing all three seasons compared to a single season for each sport. Additional studies should include more seasons.

Third, the assessment was not conducted among key stakeholders and decision-makers in the NFL, MLB, and MLS.

Future research may seek to include more data, based on the findings of this research the three seasons analyses showed an increase in predictor significance compared to the one-season analyses. In addition, future work may examine the effectiveness of special case games and their impact on HFA. For example, teams play in another country like some of NFL games when they are played in the UK or Mexico. Another feature to consider is adding more variables like weather condition and game playing time, to see how they will impact the other variables. Finally, it would be interesting to compare professional sports in other countries where the travel distance and time zone differences are not relevant.