Reducing Cost Overruns in Construction Projects Using Lean Methodology

by Sam Momtazi

B.S. in Aeronautics, December 2015, Embry-Riddle Aeronautical University
M.S. in Space Systems Engineering, May 2018, Johns Hopkins University

A Praxis submitted to

The Faculty of
The School of Engineering and Applied Science
of The George Washington University
in partial fulfillment of the requirements
for the degree of Doctor of Engineering

May 16, 2021

Praxis directed by

Timothy Blackburn
Professorial Lecturer of Engineering Management and Systems Engineering

The School of Engineering and Applied Science of The George Washington University certifies that Sam Momtazi has passed the Final Examination for the degree of Doctor of Engineering as of March 29, 2021. This is the final and approved form of the Praxis.

Reducing Cost Overruns in Construction Projects Using Lean Methodology

Sam Momtazi

Praxis Research Committee:

Amir Etemadi, Associate Professor of Engineering and Applied Science, Praxis Director

Timothy Blackburn, Professorial Lecturer of Engineering Management and Systems Engineering, Praxis Co-director

Oluwatomi Adetunji, Professor Lecturer of Engineering Management and Systems Engineering, Committee Member

© Copyright 2020 by Sam Momtazi
All rights reserved

Dedication

Most of all, my work is dedicated to a man who always believed in me. Dad, I know you are watching over me, and you are proud of what I have accomplished. I also dedicate this to my mother, my sisters, my brother, and in-laws. Thank you for your support throughout this journey.

Acknowledgements

In the words of Dr. Wayne Dyer:

“Appreciate everything and everyone. Look upon every experience you've ever had and everyone who's ever played any role in your life as having been sent to you for your benefit. In this universe, which was created by a divine, organizing intelligence, there are simply no accidents.”

First and foremost, I would like to thank my academic advisor, Dr. Timothy Blackburn, to continually support, advise, and all feedback throughout the research process and transform my thinking during this incredible journey.

I would like to thank my faculty advisors, Dr. Amir Etemadi, Dr. J.P. Blackford, Dr. Shahram Sarkani, Dr. Shahryar Sarkani, and Dr. Thomas Mazzuchi, for their helpful insights, guidance, and support, during all aspects of this research effort.

Last but not least, I would like to thank the George Washington University for providing me this amazing opportunity to pursue my doctorate.

Abstract of Praxis

Reducing Cost Overruns in Construction Projects Using Lean Methodology

Cost overrun has emerged as one of the most frequently reported problems facing projects in the construction industry. The problem of cost overruns is experienced more in publicly funded projects, such as New York City capital projects. Traditional cost control measures have failed to solve the problem of cost overruns in construction projects.

This research includes applying nine lean construction principles: supply, people involvement, planning and scheduling, quality, standardization, transparency, waste elimination, continuous improvement, and customer focus. Throughout the construction process, these principles will reduce the element of waste and non-value adding activities, and they will be applied through twenty-five selected sub-principles, which are specific techniques that reduce various forms of waste in the construction process. To investigate the effectiveness of the above lean construction principles in controlling project cost, the research adopted a quantitative approach executed through an online survey involving the managers of the various projects listed in the New York City projects database and nationwide. The study results will show which Lean Construction Principles (LCPs) critical factors will significantly reduce average cost overruns in construction projects. This research proposes a practical model called Lean Principles in Construction Management (LPCM) that can help engineering managers improve overall project performance by reducing major waste factors.

Table of Contents

Dedication ..... **iv**

Acknowledgements ..... **v**

Abstract of Praxis ..... **vi**

List of Figures ..... **xi**

List of Tables ..... **xii**

List of Equations ..... **xiii**

List of Symbols ..... **xiv**

List of Acronyms ..... **xv**

Glossary ..... **xvi**

Chapter 1—Introduction 1

1.1 Background 1

1.2 Research Motivation 3

1.3 Problem Statement 4

1.4 Thesis Statement 4

1.5 Research Objectives 5

1.6 Research Questions and Hypotheses 5

1.7 Scope of Research 6

1.8 Research Limitations 6

1.9 Organization of Praxis 7

Chapter 2—Literature Review 9

2.1 Introduction 9

2.2 History of Lean and its Transition to Construction 9

2.3 Lean Construction Principles and Sub-Principles 10
2.3.1 Supply 10
2.3.2 Quality 13
2.3.3 Planning and Scheduling 15
2.3.4 People Involvement 16
2.3.5 Continuous Improvement 18
2.3.6 Customer Focus 20
2.3.7 Waste Elimination 21
2.3.8 Standardization 26
2.3.9 Transparency 29
Chapter 3—Methodology 33
3.1 Introduction 33
3.2 Research Design 34
3.3 Research Methods 35
3.4 Target Population, Sampling Techniques, and Sample Size 37
3.4.1 Target Population and Data Source 37
3.4.2 Sampling Techniques 38
3.4.3 Sample Size 39
3.5 Research Process and Data Collection 40
3.6 Research Instruments 41
3.7 Item Analysis by Applying Cronbach's Alpha 42
3.7.1 Cronbach's Alpha 43
3.8 Data Analyses 43

3.8.1 Paired T-Test 45
3.9 Ethical Considerations 45
Chapter 4—Results 47
4.1 Introduction 47
4.2 General Information 47
4.2.1 Project Type 47
4.2.2 The Region of Project Managers' Last Project 49
4.3 Analysis of Responses Based on the Likert Scale 50
4.4 Summery of the Frequencies of the Responses 55
4.5 Item Analysis 56
4.6 Statistical Methods 58
4.7 Building the Random Forest Model 59
4.8 Partial Least Squares (PLS) 64
4.9 Partial Least Squares (PLS) Model 64
4.10 Considerations Regarding Research Hypothesis I 65
4.10.1 The Model 66
4.10.2 Final PLS Model 72
4.10.3 The Model Validation 72
4.11 Considerations Regarding Research Hypothesis II 74
4.12 Summery 76
4.13 The Research Product (LPCM Model) 78
Chapter 5—Discussion and Conclusions 79
5.1 Discussion 79

5.2 Conclusions 80

5.3 Limitations of the Study 82

5.4 Contributions to Body of Knowledge 83

5.5 Recommendations for Future Research 83

References 85

Appendix A 103

Appendix B 120

List of Figures

Figure 3.1: Lean Principles in Construction Management (LPCM)……34

Figure 4.1 Type of Most Recent Projects Managed by Construction Project Managers..49

Figure 4.2 The Region of Construction Project Managers' Last Project ……50

Figure 4.3 Repeated Cross-Validation Results.…….…..60

Figure 4.4 Relative Variable Importance Plot …….…..61

Figure 4.5 Correlation Matrix for Predictors.……...…… 63

Figure 4.6 Model Selection Plot. ……..…68

Figure 4.7 Residual versus Leverage Plot……...…….…...……..68

Figure 4.8 Distance Plot …….……...…….…….69

Figure 4.9 Score Plot…. …….……..69

Figure 4.10 Response Plot.……....70

Figure 4.11 Loading Plot ….…….……...70

Figure 4.12 Coefficient Plot …….……71

Figure 4.13 Lean Principles in Construction Management (LPCM)….……...78

List of Tables

Table 2.1 Lean Principles and Sub-Principles in Construction…….……....31

Table 4.1 The Region of Construction Project Managers' Last Project ……...50

Table 4.2 Summary of the Frequencies of the Responses…….55

Table 4.3 Cronbach's Alpha Values …….……...… 57

Table 4.4 Relative Variable Importance Ranking…...…….……..... 62

Table 4.5 Analysis of Variance ……....…….66

Table 4.6 Model Selection and Validation ……...66

Table 4.7 Model Coefficients….…….…….…….…....67

Table 4.8 Sub-Principles which are Negatively Associated with the Cost Over-run. …..71

Table 4.9 Final PLS Model……...…….72

Table 4.10 Two-Sample T-Test and CI…….....73

Table 4.11 Two- Sample Equivalence Test……. …… 74

Table 4.12 Paired T-Test.…….…..75

Table 4.13 Two-Sample T-Test for Comparing Means.…….…….……...…..75

Table 5.1 Ranking of the Most Significant Factors.…….…….80

List of Equations

Equation 3.7.1 Cronbach's Alpha (1) ….…….43

Equation 3.8.1 Paired T-Test (2)…….…….45

List of Symbols

$\sum$	Sum of all values in the range of series
μ	Mean of population values
S^2	Population samples variance estimator
Var(x)	Variance of random variable x
Std(x)	Standard deviation of random variable x
Corr (x,y)	Correlation of random variables x and y
R^2	Represents the proportion of the variance for a dependent variable

List of Acronyms

LCPs	Lean Construction Principles
TPS	Toyota Production System
LPCM	Lean Principles in Construction Management
SCM	Supply Chain Management
JIT	Just- in -Time
TQM	Total Quality Management
PPC	Percent Plane Complete
VSM	Value Stream Mapping
BIM	Building Information Modeling
PM	Project Manager
SOP	Standard Operating Procedures
TQC	Total Quality Control
RII	Relative Importance Index
PMP	Project Management Plan
TQC	Total Quality Control
VIF	Variance Inflation Factor
RF	Random Forest
MSE	Mean Square Error
PLS	Partial Least Squares
CI	Confidence Intervals

Glossary

Capital Project: A project that maintains or improves a city asset, also called infrastructure.

Schedule Changes: Period the project is behind or ahead of schedule since the last disclosing timeframe. Negative numbers represent months ahead of schedule.

Budget Changes: Number of dollars(fund) the project is under or over budget since the last disclosing timeframe. Negative numbers represent dollars under budget.

Current Phase: The stage the project is currently in as of the reporting period.

Design Phase: The design phase is when design work is being performed, resulting in drawings and specifications for the construction bid.

Construction Procurement Phase: When the procurement activities are underway, prepare the construction bid and award the construction contract.

Construction Phase: When the contractor performs work as defined in the construction contract.

Chapter 1—Introduction

1.1 Background

The construction industry has been faced with problems of inadequate safety, waste, efficiency, and effectiveness (Tezel & Niezel, 2013). Today's construction industry is characterized by increasing cost overruns, high levels of uncertainty, and variability of most of its processes. As a result, numerous measures, such as industrialization techniques, use of modern and improved computer technology, and even applying automation and robotics have been used to reduce construction delays and cost overruns (Bajjou and Chafi, 2018). However, these measures have not yet achieved a significant reduction in delays and cost overruns. According to Johansen and Walter (2007), there is a lack of an innovative strategy that can revitalize the traditional project management techniques. Studies have indicated that these solutions on their own, in the absence of an effective and responsive project management basis, cannot revolutionize the construction industry (Aziz and Hafez, 2013).

Waste has been identified as one of the critical causes of cost overruns in construction projects. Studies have indicated that waste factors can increase a project's total cost by more than 10% (Khanh and Kim, 2014; Babalola, Ibem, and Ezema, 2019). Therefore, the assessment of waste has become one of the most critical strategies to employ to improve construction projects' performance. However, studies indicate that construction project managers still conceptualize waste in terms of physical construction waste and fail to consider unnoticeable wastes such as wait time, unused employee

creativity, non-value adding activities, etc., hidden in construction processes (Emuze & Saurin, 2015).

To improve productivity and sustainability in the construction industry, a large number of studies have indicated that stakeholders in this industry are exploring new project construction processes and techniques, as well as management practices that are exhibiting the great potential to drive the construction industry towards becoming environmentally friendly and less wasteful (Ajayi et al., 2017). One such management principle and practice is lean construction, which seeks to minimize waste and maximize value to the client.

Lean construction is a management philosophy that aims to create value for the client through a systematic, rigorous approach to continuous improvement using collaborative management tools adaptive to the construction industry's characteristics and constraints (Haddad, Zouein, Salem, and Otayek, 2016). It is a practical management approach that helps to improve project productivity through the reduction of waste as well as the acceleration of the construction process (Minami, Soto, & Rhodes, 2010).

Lean construction takes its roots from the automotive industry and manufacturing, specifically the Toyota manufacturing philosophy, which was initially referred to as the Toyota Production System (TPS). Developed by the Japanese after World War II, it helped the company achieve great success in manufacturing automotive products within the stipulated time, cost, and quality (Bajjou and Chafi, 2018). Therefore, lean construction adapts the automotive industry's manufacturing philosophy to eliminate wastage and keep inventory and works-in-progress at optimum (Monde, 1998). The term

Lean Construction was first introduced in 1993 by the International Group for Lean Construction (Gleeson et al., 2007).

It is described as an approach to project design and construction with organized activities to reduce project materials waste, reduce time overruns, and direct customer value efforts. In the past few decades, construction management literature has remarkably featured lean construction concepts, but the construction industry is in the beginning stage to apply them (Ajayi et al., 2017). This represents a transition from the long-established approach to identifying what is responsible for waste and encourages avoidance instead of reacting or trying to eliminate the impacts of loss in the process of construction (Lapinski, Horman, and Riley, 2006).

1.2 Research Motivation

Various studies have pointed to the construction industry's low performance, characterized by time and cost overruns (Tiezel and Nielsen, 2013; Bajjou and Chafi, 2018). The problem is even more persistent in publicly funded construction projects such as New York City capital projects, of which reports indicate that over 47% of the city's capital projects have exceeded budget (Office of Council Member Brad Lander, 2019 https://council.nyc.gov/brad-lander/nyc-capital-project-reform/). Similarly, numerous studies have been carried out, identifying possible causes of construction project delays and cost overruns in various construction projects worldwide. Consequently, numerous studies have proposed various methods, tools, and techniques that can be used to reduce time and cost overruns (Bajjou and Chafi, 2018). However, the problem of cost overruns

and delays has not significantly reduced, even with new projects showing significant time and cost overruns.

This poses a challenge on one hand while, on the other, it motivates to seek more innovative and effective ways of solving the problem of cost overruns in the construction industry.

This study's motivation is to develop a lean model capable of helping engineering managers improve general project management by reducing key waste factors in the construction process.

.

1.3 Problem Statement

On average, 47% of New York City Capital Projects over $25 million have exceeded budget by an average of 12% in the last seven years. (Statistics of Office of Council Member NYC, 2019).

1.4 Thesis Statement

By applying Lean Construction Principles (LCPs), New York City construction projects' cost overruns will reduce significantly.

1.5 Research Objectives

The main research objectives are:

I. To determine the effect of Lean Construction Principles on construction. This research seeks to establish whether lean construction principles can reduce cost overruns in construction projects. Being a quantitative study, the analysis will estimate the percentage by which the applied lean construction principles reduce cost overruns in capital projects.

II. To identify significant causes of cost overruns in capital projects.

III. To propose a lean construction model that can help engineering managers improve overall project performance by reducing major waste factors in the construction process.

1.6 Research Questions and Hypotheses

RQ1: What are the key causes of cost overruns in construction projects?

RQ2: What is the effect of using Lean construction principles on construction projects experiencing cost overruns?

H1: Projects that utilize Lean Construction Principles (Including Waste Elimination, Supply and People Involvement, Planning and Scheduling, Standardization, Quality, Transparency, Continuous Improvement, and Customer Focus) tend to have less cost overruns.

H2: The application of Lean Construction Principles (LCPs) will significantly reduce cost overruns in New York City construction projects.

1.7 Scope of Research

The research was carried out by conducting an online survey of the project managers in New York City and nationwide. Emails were sent to notify the managers about the research; those with responses were included in the study. The survey involved construction projects both in the private and public sectors.

The survey was completed by NYC senior supervisors and project managers who worked for the city and were directly involved in controlling construction activities in the respective construction projects. The survey also involved other capital projects in various cities in the U.S. to provide a variable population sample that could boost the study results' variability. The survey questionnaire was simplified, such that even project managers who were not familiar with lean construction could complete it.

1.8 Research Limitations

Most of the limitations of this study result from the research design. Having taken a quantitative model, the research only focuses on determining the relationship between the study variables, i.e., The project cost and lean construction principles fail to ascertain deeper underlying meanings about the possible causes of cost overruns in construction projects (Denzin and Lincoln, 1998).

The positivism applied in a quantitative study cannot explain how people interpret their actions and others (Blaikie, 2007). A quantitative study involves a structured questionnaire with closed-ended questions, limiting responses and outcomes to the proposal listed. Apart from limiting the participants' responses to the study's options, the

results might fail to capture the construction projects' actual situation as the study is undertaken in a generalized form (Saunders, Lewis, and Thornhill, 2009).

1.9 Organization of Praxis

The research will have five chapters organized as follows:

Chapter 1 will present the introduction, which provides the background information, the motivation for carrying out the research, and its objectives. The introduction also defines the study's scope, indicating that the study will involve an online survey targeting the project managers nationwide. This chapter also gives the limitations of quantitative research and ends with a brief overview of the study's organization.

Chapter 2 will provide a literature review of the relevant literature about the topic. The review will focus on the causes of cost overruns, give a detailed analysis of the lean construction principles available in the literature, and touch on the interactions that will form a new construction model proposed by this study. The literature review will also provide both the theoretical and conceptual models of the study.

Chapter 3 will present the methodology, a detailed explanation of the quantitative research method used in the study. The methodology will explain the questionnaire's developmental process, the data analysis process using the *Random Forest* machine learning and *Partial Least Squares* (PLS). The methodology will highlight the challenges encountered in the research process and end with the ethical considerations involved in the study.

Chapter 4 will present the study's findings and provide detailed discussions on the current research results about the previous conclusions recorded in the literature review.

Chapter 5 will present the study's conclusions, determining the research question, i.e., whether lean construction principles can reduce cost overruns in construction projects. This chapter will also recommend a new lean construction model that engineers can use to improve construction projects' performance.

Chapter 2—Literature Review

2.1 Introduction

The chapter presents a critical review of various types of lean construction principles and practices that have been developed in different countries and used to reduce cost overruns in the construction industry. The review goes further into the subprinciples proposed in various lean construction models and seeks to establish how they have been used to mitigate the causes of waste in the construction process.

2.2 History of Lean and its Transition to Construction

The *lean manufacturing* concept originated from the Toyota Production System (TPS), referred to as the lean manufacturing philosophy after the Second World War, to enhance products' production and consumption efficiency in the automotive industry. However, the concept of *lean* dates back to the reproducible activities of Fredrick Taylor, who is known as the Father of the scientific management & efficiency movement, referred to as Taylor's Theory. Its implementation was achieved through the conveyor belt invented by Henry Ford that facilitated mass production in the 19th century. However, the concept of mass manufacturing experienced a shift in 1949 when sales for Toyota declined, leading to a significant portion of its workforce's retrenchment. The company revised its production, leading to the Toyota Production System (TPS) that marked a shift from Henry Taylor's mass production to lean production in the 1990s (Sarhan, Xia, Fawzia, and Karim, 2017).

After that, the Toyota Production System that made the concept of lean production was applied along with Total Quality Control (TQC) to reduce waste and defects in the manufacturing process. The lean production concept was later adopted in Western countries, where it was introduced as lean thinking. Although it was initially used in manufacturing, the concept of lean is increasingly being adopted in the construction industry in many countries, where it is referred to as *Lean Construction* (Bajjou & Chafi, 2018).

2.3 Lean Construction Principles and Sub-Principles

This section identifies the key principles and sub-principles developed in various countries to mitigate waste and the more significant problem of cost overruns in the construction process through a detailed review of the relevant literature. The section also illustrates how key lean construction principles are applied to mitigate waste sources in the construction process.

2.3.1 Supply

This refers to an efficient logistics management system that ensures the availability of all the necessary materials, tools, and labor to execute the construction process successfully. The supply principle seeks to establish an optimal flow in the construction company’s logistics to reduce stock levels to only what is necessary at a given time in the construction process (Dupin, 2014). Studies have pointed out that the construction process's success can be achieved by organizing materials and equipment in the construction site. This principle can be achieved through four significant sub-

principles, including supply chain management, pull system, just-in-time, and suppliers' involvement. The subprinciples are described as follows:

I. Supply Chain Management (SCM)

Supply Chain Management (SCM) is an innovative and revolutionary approach to project management that involves integrating and synchronizing inter-dependent players in the construction cycle. It is more than planning material and information flow in the project; it entails a joint effort to achieve value by optimizing resources, core competencies, information, and relevant technologies to achieve construction project goals. SCM involves managing internal and external resources to ensure a smooth flow of construction activities (Aneesa, Gupta, and Desa, 2015).

II. Pull System

Theorists and manufacturing management experts have distinguished between two major approaches to managing workflow, i.e., push and pull systems. As illustrated in most traditional project management techniques such as Critical Path Method (CPM) and Material Requirement Planning (MRP) systems, Push-based approaches, as illustrated in most traditional project management techniques, release work into the production process based on pre-determined criteria such as delivery dates. On the other hand, the new pull-based thinking for engineering and manufacturing advocates for pull-based workflow systems that work into the production process based on its state. For instance, a simple pulling technique limits the amount of inventory left between construction worksites, so materials are allowed to move forward to the next construction phase only when their inventory space is empty (Ballard, 1999). The pull system relies on

a cascading delivery mode from downstream to upstream phases of the construction process such that nothing can be produced in the upstream unless there is a need demonstrated from the downstream construction stages (Diekmann et al., 2004).

III. Just-In-Time (JIT)

Just-in-Time (JIT) is a management philosophy that aims to provide the required quantity of material at the right place in an appropriate order at the right time (Gudehus and Kotzab, 2009). JIT is a concept based on a pull system requiring that materials are delivered when and where they are required for the construction process. Just-in-time advocates for frequent deliveries on time, minimizing on-site storage (Harker, Allcorn, and Taylor, 2007). JIT is applied in the construction industry to increase construction processes' efficiency and productivity (Low Suit and Gao, 2011).

IV. Suppliers' Involvement

Studies indicate that contractors spend up to 90% of their project turnover in purchases of materials, equipment, and other services (Vrijhoef and Koskela, 2000). However, studies also indicate that contractors have not taken the initiative to build contractor-supplier cooperation because construction projects are temporary, and this poses significant difficulty in managing long-term supplier-contractor relationships, which causes a great deal of dissatisfaction (Briscoe and Dainty, 2005).

There is a need to develop and nurture long-standing contractor-supplier relationships to reduce material delivery delays to construction sites (Tezel, 2007). Effective management of contractor-supplier relationships requires optimizing the supplier bases, enhancing the existing purchasing processes, setting the performance

standards higher, and matching them with effective supplier development strategies (Bemelmans et al., 2012).

2.3.2 Quality

Studies have indicated that quality defects leading to re-work increase construction projects' costs by over 52% and significantly reduce customer satisfaction (Forcada et al., 2017). However, studies have also found that introducing quality in the design, material purchase, and the actual construction process can improve the finished construction project. Simultaneously, research has established that adopting a quality management system that ensures all construction activities and processes are under control will help quickly identify the root causes of quality defects in the construction process (Mallawaarachchi and Senaratne, 2015). Various studies have proposed different techniques in ensuring quality is maintained in the construction process. In their study, Bajjou and Chafi (2018) identify total quality management, poka-yoke devices, and the five why's as some of the key lean techniques that could identify the root causes of quality defects in the construction process. They are discussed as follows:

I. Total Quality Management (TQM)

Quality in the construction industry refers not only to the quality of the finished construction product; it speaks to the total management approach towards completing the construction project. Total Quality Management (TQM) is a management philosophy that identifies quality as the driving force behind the design, planning, and improvement in project activities and processes (Adedeji, 2018). In the construction industry, TQM refers

to a management approach that involves all organizational functions, including engineering, construction, and customer service, that is directed towards meeting the project objectives, emphasizing fulfilling clients' needs (Ullah et al., 2017).

II. Error-proofing (Poka-Yoke)

Poka-Yoke (error-proofing) is a Japanese strategy for mistake-proofing to prevent defects or nonconformities in the manufacturing process. Poka-Yoke is used to identify and eliminate the causes of variation in the manufacturing process (Rajan, Dwivedi, and Ajay 2016). In the construction industry, error-proofing involves using automatic devices and modern technology to make it difficult for an error to occur or to detect an error immediately when it occurs. This is a lean method of error or mistake management that acknowledges mistakes, and it seeks to reduce them throughout the construction process (Tommelein and Demirkesen, 2018). By reducing quality defects, Poka-Yoke contributes significantly to reducing project costs associated with construction errors (Diekmann et al., 2003).

III. The Five Whys

The five whys technique is a problem-solving technique that begins with the result, facilitates reflection on the possible cause of the problem identified, questions, and answers five times. This is a simple problem-solving technique that promotes deep thinking by questioning the proposed solutions' construction process. “There are three critical elements of effective use of the five whys technique: accurate, complete identification and statement of the problem, honesty in answering the five questions, and an attitude of determination to reveal the cause and solve it” (Serrat, 2009).

2.3.3 Planning and Scheduling

In the construction industry, planning is a fundamental but challenging task in executing and managing projects. Planning involves choices such as the technology to be used, the definition of scope and tasks, the nature and magnitude of the resources involved, and the organization and the technical aspects of the construction project. While developing plans for construction projects, project managers use schedules to control the construction project's activities. Planning and scheduling advocates for proper precedence of activities and maintenance of efficient schedules sensitive to available alternatives and an inherent assumption that ensures low project costs are involved and projects are completed within a stipulated or reasonable time (Sacks, Treckmann, and Rozenfeld, 2009).

In the construction industry, scheduling needs to consider the trade-offs between time and cost such that it minimizes the consumption of project resources (Yang, 2007) while minimizing project time (Elmaghraby, Herroelen, and Leus, 2003). Similarly, in the construction industry, planning for project resources should lead to the development of reliable projects (Kerzner, 2009). Various studies have identified various planning and scheduling techniques. In this regard, in the study, Bajjou and Chafi (2018) identified collaborative planning and the Percent Plan Complete (PPC) as key lean principles in the construction project management process; they are discussed as follows:

I. Percent Plan Complete (PPC)

This a measure of how well the planning systems of a construction project are working. The Percent Plan Complete (PPC) indicator is used to track a construction process's accomplishments determined by the number of promises or the number of

activities completed at any given day divided by the number of activities planned for that week. The PPC indicator measures the rate of implementation of various project commitments in a given duration. For instance, PPC can establish the percentage of completed construction activities among scheduled ones (Nesensohn, Demir, and Bryde, 2018).

II. Collaborative Planning

Collaboration has been identified as a crucial factor in the design and execution of construction projects. Studies have pointed out that traditional project management has proved insufficient in managing projects in the modern engineering and construction project management environment, becoming increasingly complex and unpredictable (Mitchell et al., 2011). Research has demonstrated that collaborative planning. When augmented with other project management tools such as process mapping, it can increase project predictability among other construction project performance metrics. As an act of bringing people together to make plans for their work activities, it is essential to note that collaboratively planning is not an end to itself but a catalyst for other project management processes (Yin, Qin, and Holland, 2011).

2.3.4 People Involvement

The principles of lean construction consider the workforce as the significant factor that improves and changes the firm. Integrating top management and the workforce empowers them better and ensures their full engagement in the firm's attempt to improve (Bajjou & McCaffer, 2013). Individual initiative is encouraged when it comes to a lean

construction firm. Hence, the project's productivity turns to the support of every stakeholder. Such collaboration involves developing human resources via individual centralizing development (Bajjou & Chafi, 2018).

The construction industry's significant characteristics include the lack and shortage of a skilled workforce, which presents a significant hurdle influencing the successful implementation of lean construction. Other challenges that could arise due to unskilled and poor-quality workers include the workforce's low productivity levels, errors in construction, and quality defects, thereby affecting the project's general performance (Bajjou et al., 2019). "The main sub-principles of people involvement include workforce involvement and teamwork" (Bajjou et al., 2019).

I. Workers' Involvement

Workers are the essential resources of the construction company's performance, project success, and customer satisfaction. Workforce involvement at every level of the decision-making process, including contribution and capitalization of knowledge, is essential in the construction industry. Workforce shortages are widely experienced in this industry, leading to waste in terms of time and production. The construction process is improved when employees are engaged fully in the decision-making process and have a high-quality work-life (Nesensohn et al., 2015). Throughout the organizational process, a workforce feels empowered and becomes more productive since they know they are part of the decision-making process. Workplace involvement is also achieved through employee motivation. Construction projects require workers who are actively searching for ways to enhance work and initiate performance improvements. Construction work is

improved when employees' successes and efforts are acknowledged and rewarded, resulting in increased employee motivation (Aziz & Hafez, 2013).

II. Teamwork

There is collaborative planning involving the designers, suppliers, sub-contractor, and the primary contractor when it comes to lean construction philosophy. Teamwork makes the workplace safer and more efficient for everyone at the construction site. Effective teamwork includes open communication, and every worker is allowed to contribute to improving the job site to successfully and safely achieve the construction goal (Salem et al., 2006). Teamwork begins with pre-planning and organization. Setting clear expectations of every team member guides them to work towards a common goal of improving quality. Excellent communication means that every worker is allowed to give his or her opinion, which helps build trust within workers. Effective teams are built through trust and the sharing of information. Construction is a complex activity; thus, when individual workers commit to accommodate each other to build a team, injuries are minimized, and productivity and quality are improved (Tezel & Nielsen, 2013).

2.3.5 Continuous Improvement

Many researchers have stated that the lean construction process is not a destination but a journey (Anvari et al., 2011; Khanh & Kim, 2015). Hence, the permanent renovation of conventional management strategies and improved performance through a systematic approach that seeks excellence towards resisting the construction market's competitiveness and pressures becomes necessary for construction firms

(Bajjou, Chafi & Ennadi, 2019). Construction firms need to be relatively stable towards benefitting from continuous improvement. The firm's practices are ascertained by the leadership of the contractor's organization to facilitate continuous improvement.

Strategic partners like subcontractors, designers, and suppliers are involved in the learning cycle and the root causes analysis (Bajjou et al., 2019). "The primary sub-principles of continuous improvement are metrics (productivity, quality, safety) and huddle meetings" (Bajjou et al., 2019).

I. Metrics (Productivity, Quality, and Safety)

The focus here is measurable and actionable improvement instead of developing capabilities. The indicators of performance, such as inventory, cycle time, and quality defects, are measured, recorded, and analyzed in a clear, objective, and systematic manner in order to drive the project towards meeting customer needs and improving the construction performance process (Devaki & Jayanthi, 2014). These metrics include defining the different flow processes and concentrating on the bottlenecks to smooth out and speed up material and information flows. Measures are essential in the pursuit of lean construction. Measures such as productivity, quality, and safety offer access to continuous improvement through pinpointing improvement potential and checking progress achieved (Tezel & Nielson, 2013).

II. Huddle Meetings

Two-way communication is fundamental to the daily huddle process in order to accomplish worker involvement. With some training and problem-solving involvement and increased project awareness, employee satisfaction increases through improved self-

esteem, job meaningfulness, and a sense of growth (Salem et al., 2005). Huddle meeting techniques aim to ensure workforce involvement, rapid response to problems, and continuous open communication through daily start-up meetings (Jorgensen & Emmitt, 2008).

2.3.6 Customer Focus

Customer focus principles aim to introduce the customer to lean construction by identifying the necessary value and ensuring that its flexibility is maintained (Bajjou & Chafi, 2018). Further, customer focus helps the contractor and customer have a great relationship, reducing waste and maximizing value throughout the task's execution (Issa, 2013).

I. Flexible Resources

Construction firms need to respond quickly to the current market conditions and the clients' needs (Diekmann et al., 2004). Construction companies need to use flexible resources and adopt planning systems, including getting required equipment and materials immediately and speedily adapting workers towards meeting what new customers require (Diekmann et al., 2003).

II. Optimize Value

There are analysis and integration of what customers require into the identification of the added value. These needs determine the required construction (Mossman, 2009). This sub-principle helps to eliminate waste in a process that does not create value for customers. Most of the issues encountered by construction companies

include eliminating waste in a sub-process rather than the whole process, which ends up optimizing that sub-process instead of optimizing the project as a whole. Waste is defined based on the customer's value requirement. Construction processes are executed based on customers' needs (Cui et al., 2014).

III. Customer Relationships

It is possible to achieve a high-level relationship with customers by increasing interaction between the construction firm and its clients by assessing risk, dialogue, and transparency (Bygballe & Ingemansson, 2014). While demonstrating each construction process, there is a focus on cooperating value creation with the user. The client and contractor need to work together to eliminate what is not valuable to the customer and not enhance the performance process.

A productive customer relationship ensures that the contractor engages the customer throughout the construction process, and he can communicate his needs early before construction starts. Both the contractor and the customer monitor the construction process, and in case there is any deviation, it is immediately corrected to create value for the customer (Savolainen et al., 2015).

2.3.7 Waste Elimination

Waste elimination is the core of lean construction philosophy. This dimension mainly aims to integrate the culture of eliminating all types of waste, such as over-production, inventory, waiting, defects, displacements, unused employee creativity, unnecessary transportation, and over-processing (Khanh & Kim, 2014). Lean

construction is the most effective method to minimize or eliminate in the construction industry. Waste is the share of activities that do not add value to the construction process, company, or customer. Reducing or eliminating activities that do not add value is one of the core strategies for improving the construction process (Bajjou et al., 2017). The key sub-principles of waste elimination include an optimized product system, TIMWOOD, waste disposal management, and Value Stream Mapping (VSM).

I. Optimize Product System

To optimize a product system involves the placement of a construction unit output in the construction unit of its success as an input as well as within a progressive cycle, at the same time using minimum amounts and keeping up a flexible number of employees based on what their successor construction units require (Tezel & Nielsen, 2013). As they continue to engage output as an input while constructing, there is the maintenance of a remarkable and progressive enhancement of the process performance.

Optimization of products helps to find additional value in the company's construction system. Therefore, there is a need for the construction company to optimize its production system to add value and improve the construction process (Bajjou et al., 2017).

II. TIMWOOD

TIMWOOD is an abbreviation for the seven sources of waste in construction. They include Transport, Inventory, Motion, Waiting, Over-processing, Over-production, and Defects. It is argued (El-Namrouty and Abu Shaaban 2013) that perhaps TIMWOOD is the most considerable focus by the lean construction philosophy that seeks to eliminate

all forms of waste identified through the TIMWOOD framework. TIMWOOD can help to reduce project costs, improve profit, improve project lead times, and enhance customer satisfaction (El-Namrouty and Abu Shaaban, 2013). As a waste elimination approach, TIMWOOD reduces waste in the following ways:

1. **Transportation**

Transportation refers to the movement of materials between different workstations in the construction site that does not add value to the construction process but adds time to the construction cycle, contributing to inefficiency in terms of labor and space (Bajjou and Chafi, 2018).

Unnecessary Transportation of materials can be reduced by placing time close to the point of use and simplifying material handling systems.

2. **Inventory**

Inventory as a source of waste in the construction process refers to stored on-site materials that will not be used in the short term, leading to other problems, such as storage costs, inadequate space, and degradation of the quality of materials (Bajjou and Chafi, 2018).

Excess Inventory can be eliminated by adjusting production processes to enable a smooth flow of work, working with small batch sizes, and observing the first-first-out inventory management method.

3. **Motion**

Motion as a form of waste in the construction process refers to the movements of operators and materials that do not add value to any construction unit's production cycle

but increase the difficulty of doing work. Bajjou and Chafi (2018) indicate that construction workers walk between 7 and 9 kilometers in a day due to numerous factors, such as searching for equipment, construction material, and information.

Waste due to unnecessary Motion can be reduced by improving the workstation layout to avoid unnecessary walking, bending, or reaching and arranging construction activities to facilitate smooth flow from one station to another (Tony, 2020).

4. Waiting

Waste in terms of *waiting* time results from failures in synchronizing planned tasks and lack of required material, labor, equipment failure, or any other necessary information in undertaking construction activities. In traditional project management, waiting time is one of the most common causes of waste in the construction process (Bajjou and Chafi, 2018).

Unnecessary Waiting can be solved by proper synchronization and coordination of construction activities.

5. Over-Production

Over-production refers to the production of more units than the customer requires or constructing structures that are not necessary and will be destroyed in the short term (Bajjou and Chafi, 2018).

Over-production can be solved by working with smaller batches and using accurate information about customer requirements or demand for the construction units.

6. Over-Processing

Over-processing points to tasks and activities that lead to unnecessary, extra work and more accurate, complicated, or costly tools than necessary. This increases the cost but adds little or no value to the construction process's final product (Bajjou et al., 2017).

Over-processing can be reduced by standardizing the best construction techniques and setting precise specifications and quality acceptance standards.

7. Defects

Defects contribute to waste generation in the construction process in two ways: in terms of labor, materials, and equipment that are lost in the process or the level of turnover that is missed during the entire time of repairs and other means of correcting the defects. The defects also distort the flow of activities in the synchronized or planned flow of construction activities and reduce site workers' motivation as they repeat the affected parts (Bajjou and Chafi, 2018).

Defects can be reduced by improving construction processes, reducing errors, and using high-quality construction materials (Tony, 2020). By identifying waste sources in the construction process, TIMWOOD emerges as one of the most critical sub- principles of waste elimination in the construction process.

III. Waste Disposal Management

Waste disposal management is essential in the construction industry. A new concept has been suggested called *3R,* which stands for Reuse, Recycle, and Reduce. The concept of *3R* has been suggested to assist in reducing, reusing, and recycling waste in the construction industry. All various processes involved in construction projects, such as

designing, manufacturing, and production at the construction site, relate to waste creation; thus, *3R* principles need to be applied to manage waste in the construction projects (Aadal, Rad & Fard, 2013). *Reduc*e means decreasing the extreme form of energy waste in the construction company. *Recycling* helps to use used materials, such as cycling energy wasted through air conditioners' heat, thus creating value for the construction company. *Reuse* is used to convert energy to other forms at various stages of construction activities (Udawatta et al., 2015). Waste disposal management is essential in the construction industry since it helps to reduce production costs and adds value to the construction process.

IV. Value Stream Mapping (VSM)

VSM tools help to represent the construction process graphically; hence, it becomes clearer to recognize and address the sources of waste. The construction process's productivity can be enhanced by proposing future VSM that should be achieved (Abdulmalek & Rajgopal, 2007).

VSM helps collect value-added and non-value-added actions in the construction process, which is critical towards using main flows to bring together a product or a group of products utilizing similar resources (Al-Aomar, 2012). Value stream mapping is aimed at the identification of every form of waste and its immediate removal.

2.3.8 Standardization

Standardization describes the extensive processes, methods, or components in which there is repetition, regularity, and a background of successful predictability and

practice (Wang & Ma, 2014). Standardization includes a set of analysis tools that results in a set of Standard Operating Procedures (SOP). SOP describes all operator tasks, such as control processes, cycle time, work sequences, and process steps. Standardization is an essential technique for designing an efficient way that helps to build creating techniques in the shortest time possible using minimum effort (Hook & Stehn, 2008). Construction companies need to ensure standardization is achieved through prefabrication (an innovative technology) since it is a critical technique in reducing waste. Standardization of work means that construction guidelines and processes are clearly defined and communicated to eliminate variations in how the work is performed.

I. Optimizing Work Content

The construction process is based mainly on pre-assembled, standard, and prefabricated construction elements (Trivedi & Kumar, 2014). Prefabrication and assembly of components other than the construction site help control construction costs by regulating time, wages, and materials. Optimizing work content aims at finding the best possible combination of elements such as input, products, time, capital, people, and their creations to minimize waste and, at the same time, to add value to the customer. Optimal work content minimizes the possibilities of failure. A lean construction approach uses simulation methods that determine the productive amount of work content and the least cost to achieve the highest benefit and performance of the construction process (Tiwaril, Magar & Honnutagi, 2018). Optimizing work content can lead to a substantial reduction in the cost and duration of repetitive construction projects.

II. Takt Time

Takt time is defined as the rate of production that matches the demand for the product. Thus, takt time is where consumer demand is known through the market (Kim & Park, 2006). Takt time can be deduced using the available working time and customer demand.

Defining the rate of demand is difficult for construction managers in the process of construction. This challenge is solvable by looking at the available time to finish the construction work and verifying it for demand. It is likewise solvable by examining the practicality of enhancing the slowest trade capacity as it hinders the process of construction (Frandson et al., 2014). Hence, "it is the responsibility of construction managers to improve production rate, so it moves at the same speed as takt time" (Yassine et al., 2014).

III. Work Sequence

"Work sequence offers a detailed description of every construction task as it specifies the minimum resources, cycle time, and takt time for the activity" (Diekmann et al., 2004). Specific peculiarities characterize construction processes. These processes' complexity does not allow for an overall process approach valid and can easily be adapted to each construction project. Therefore, construction managers need to develop a sequence of construction tasks to ensure completeness and logical correctness in the construction process. A lean construction modeling technique can develop a flexible work sequence that can be easily edited when needed. Work sequence defines the order in which a construction worker should perform tasks, including processes and motions.

This ensures that the workforce performs the tasks in the most similar manner possible to minimize variation and, ultimately, defects. Preparation of work helps avoid unnecessary production and eliminates time wasted during waiting (Huhnt, 2005).

2.3.9 Transparency

Transparency involves the construction process's ability to communicate or interact with people (Formoso et al., 2002). The primary purpose of having transparency in the construction process is to maintain a flexible, transparent, and straightforward workflow through the *5S* approach and visual management, which assists in improving the organization of the construction site and in ensuring communication reliability between the different partners (Sacks, Treckmann & Rozenfeld, 2009). Application of visual tools reduces the interdependence between the units of construction, allows immediate recognition of process status, maintains a clean and organized work site, makes visible the invisible qualities through indicators, integrates information into the construction process, and uses signage and layout intending to make the process directly observable (Tezel et al., 2010). The main sub-principles of transparency include workplace organization and Building Information Modeling (BIM).

I. Work Organization

Workplace organization helps to maintain a clean and organized worksite. The *5S* technique, which stands for Sort, Set in order, Shine, Standardize, and Sustain, is considered one of the most appropriate in keeping the worksite organized. It also increases the workforce's safety and reduces damage to materials or equipment caused by

improper storage (Aziz & Hafez, 2013). The workplace's low arrangement can cause injuries to the workers, increase retrieval time, damage products, and result in a lack of workspace. These issues undermine the performance of the construction process, making it ineffective. Construction managers need to organize the workplace to improve the construction process and eliminate waste creation due to an unorganized workplace (Ikuma, Nahmens & James, 2011). Lean integration into the construction process can help to improve the performance of the process.

II. Building Information Modeling (BIM)

Building Information Modeling (BIM) is a crucial tool for transferring and grouping data and information between the various actors of construction projects (Al Hattab & Hamzeh, 2015). BIM significantly reduces conflicts with design by relying on one source of information and enabling clash checking. It enables a better evaluation of function and visualization of form. BIM also allows more natural creation of design alternatives, design approach integrity, and better maintenance of information, including reliance on a single source of information. BIM changes worksite processes and reduces the cycle time of drawing construction, thus eliminating waste from every step of the construction process. Furthermore, BIM adds value to the client by completing value-adding activities or functions as quickly and as effectively as possible (Tauriainen et al., 2016).

2.4 Summary and Conclusion

The chapter provides a critical overview of the key elements in lean construction. A detailed review of the existing body of literature on the concept of lean principles led

to identifying nine principles that can be applied in the construction industry to mitigate the problem of cost overruns.

These principles can be implanted in the construction process through the twenty-five sub-principles discussed throughout the literature review chapter. The principles and sub-principles are shown in table 2.1 below:

Principles	*Sub-principles*
1. Supply	1.1 Supply Chain Management (SCM) 1.2 Pull System 1.3 Just-In-Time (JIT) 1.4 Suppliers' Involvement
2. Quality	2.1 Total Quality Management (TQM) 2.2 Poka-Yoke (Error-Proofing) 2.3 The Five Whys
3. Planning and Scheduling	3.1 Percent Plan Complete (PPC) 3.2 Collaborative Planning
4. People Involvement	4.1 Workers' Involvement 4.2 Teamwork
5. Continuous Improvement	5.1 Metrics (Productivity-Quality-Safety) 5.2 Huddle Meetings
6. Customer Focus	6.1 Flexible Resources 6.2 Optimize Value 6.3 Customer Relationships
7. Waste Elimination	7.1 Optimize Product System 7.2 TIMWOOD 7.3 Waste Disposal Management 7.4 Value Stream Mapping (VSM)
8. Standardization	8.1 Optimize Work Content 8.2 Takt Time 8.3 Work Sequence
9. Transparency	9.1 Work Organization 9.2 Building Information Modeling (BIM)

Table 2.1 Lean Principles and Sub-Principles in Construction

These lean construction principles overlap in their application, allowing project managers to apply any relevant principle at any phase of the construction process to prevent any form of waste and to complete the project on budget.

Chapter 3—Methodology

3.1 Introduction

The methodology chapter presents a detailed explanation of the method used in carrying out the research. The chapter explores the philosophical underpinnings informing the choice of the research design adopted and the research methods deployed to collect and analyze data.

The chapter gives a detailed description of the research population and the sampling method used to identify the study participants, and the ethical considerations affecting the study's conduct. This research proposes a new model called Lean Principles in Construction Management (LPCM), and it includes the application of *eight* principles and *ten* lean sub-principles in construction. The principles and sub-principles are shown in figure 3.1 below:

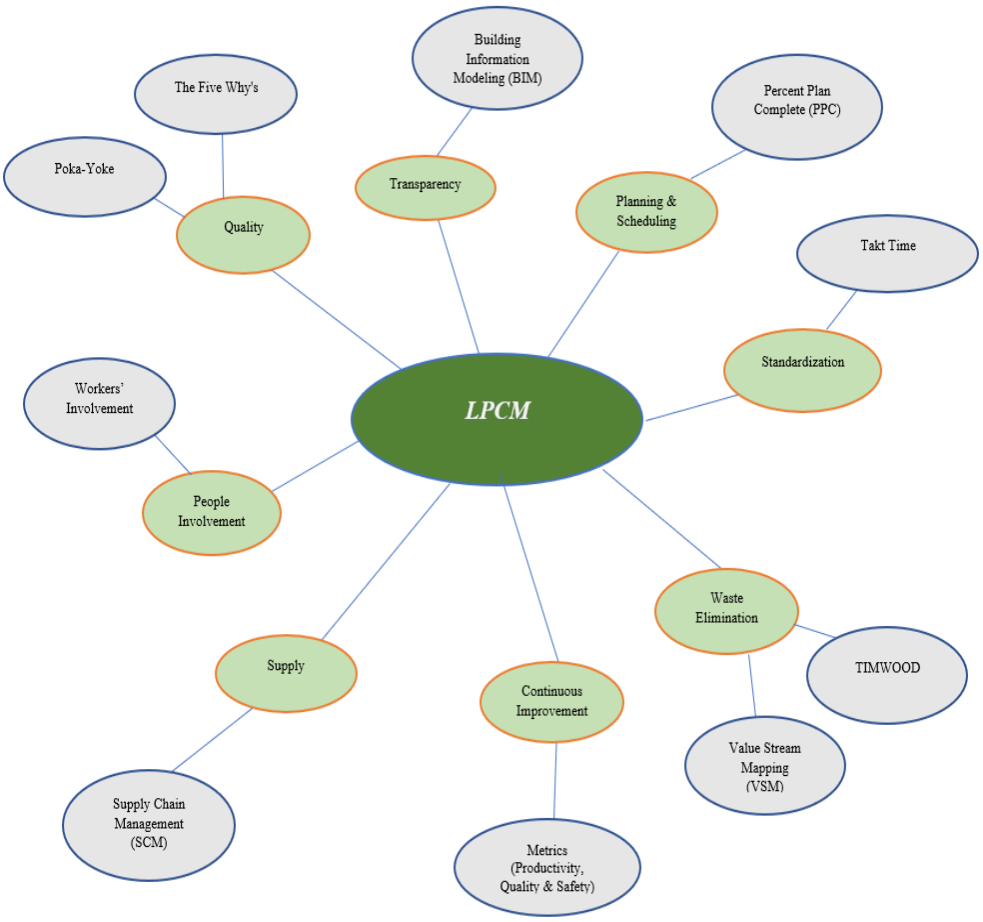


Figure 3.1: Lean Principles in Construction Management (LPCM)

3.2 Research Design

Based on the philosophical assumptions discussed above and practical considerations of the data's nature to be collected, the research adopts a quantitative research design. The quantitative research design involves quantifying and analysis of research variables to get the results of the study. A quantitative approach involves collecting numerical data and analyzing it using statistical methods (Aliaga and Gunderson, 2002).

Quantitative research design starts with identifying and explaining the research problem and proceeds to the formation of hypotheses and research questions. It entails reviewing the existing literature and collecting and analyzing quantitative data using scientifically proven instruments, yielding statistical data that is interpreted to form the research findings (Creswell, 2003; Williams, 2011). A quantitative research approach makes it possible to test the study's hypotheses using scientific methods empirically, and the findings of a quantitative study are generalizable to other populations (Johnson and Christensen, 2006). A quantitative research design is appropriate in this study owing to its ability to establish the existence, direction, and strength of the relationship between the dependent and independent variables for the study (Creswell, 2009).

3.3 Research Methods

The research was carried out using an online survey. The significant strengths of an online survey are discussed as follows:

I. The United States Outreach

One of the significant strengths of an online survey is the United States outreach. With the advent of internet technology, it is now possible to access broader populations with a wide range of characteristics using internet resources and accompanying technologies at a low cost within a short period (Evans and Mathur, 2005).

II. Flexibility

Online surveys are highly flexible. They can be carried out in various formats, including sharing a link with the research survey through the target

population's emails or inviting visitors to a website to participate in a survey. As technological advancements emerge, more innovative ways to access geographically diverse populations using internet-based technologies will emerge (Evans and Mathur, 2005).

III. Timeliness

An online survey provides a time-efficient approach to carry out research. Given all the interaction about the survey occurs over the internet. The approach minimizes the time taken to get the survey into the field to collect data and the time taken to collect the respondents' responses. The online survey involves real-time access to geographically diverse research populations. The surveys can be shared in various chat rooms, focus groups, and social media sites, which increase the speed of access of the survey with just a click of a button (Evans and Mathur, 2005).

IV. Convenience

On top of having to respond to the survey from any geographical location with internet accessibility, online surveys offer a great deal of convenience in various ways. For instance, respondents can complete online surveys at an appropriate time. Some surveys offer flexibility in that respondents can resume responding if it is stopped due to any reason. As opposed to telephone surveys where a participant is kept waiting, an online survey makes it possible for the respondent to respond to the survey conveniently (Hogg, 2003).

V. Ease of Data Entry and Analysis

An online survey is a simple, straightforward way to complete a survey, making it less challenging to tabulate and analyze the data. Besides saving the administrative struggles to send the questionnaires to the target population and collect responses from the field, the data collected through online surveys is instantaneously stored in a database, ready for preparation and analysis (Wilson and Leakey, 2003).

VI. Question Diversity

Online surveys are capable of holding different types of questions. For instance, online surveys can comprise multiple-choice questions, multimedia formats, scales, or even open-ended questions. Online surveys are capable of accommodating many types of questions (Evans and Mathur, 2005).

3.4 Target Population, Sampling Techniques, and Sample Size

3.4.1 Target Population and Data Source

The target population refers to the area of interest in the research. It can also be referred to as the population of interest. Research protocol demands that the target population's demographic characteristics, such as gender, age, socioeconomic, or work status, are described so that an ideal participant is identified (Majid, 2018). In a quantitative study that requires numerical data collection, the target population is viewed as the data source (Majid, 2018).

3.4.2 Sampling Techniques

In research, it is not appropriate or feasible to study the entire target population. In a quantitative survey, the population sample selected ought to represent the population sample to generalize the study's findings to its entire population (Van den, Sandoy, and Brestoff, 2013).

Although the research has taken a quantitative approach, which requires a probabilistic sampling technique such that every item in the target population has an equal chance of being included in the research sample as indicated in (Taherdoost 2016), practical considerations, mainly the ability to have first-hand information about the cost of the construction project identified and to make decisions that have a cost implication on the construction project led to the adoption of a non-probabilistic sampling technique, particularly purposive sampling.

Purposive or judgmental sampling is a non-probabilistic sampling technique in which the research participants are selected deliberately to provide information that may not be readily available from other sources (Taherdoost,2016). Purposive sampling entails selecting potential respondents who are perceived to have knowledge or experience about the problem being addressed by the study.

In their study on sampling methods, Creswell and Plano (2011) observed that purposive sampling makes it possible to reach respondents who are both willing and available to participate in the study. Another study (Taheroost,2016) indicated purposive sampling is a low-cost, convenient, and time-saving research technique; however, the author quickly notes that purpose sampling is highly subjective, making it inappropriate to generalize the findings in the research.

When using purposive sampling, the respondents' selection of the study should be carried out using established eligibility criteria (Taherdoost,2016). In this study, the inclusion criteria involved managers involved in managing and supervising construction projects in both New York City capital projects and other construction projects throughout the United States. In a few instances where the construction project manager could not be reached, the immediate supervisor could respond to the survey. In both New York City and nationwide construction projects, construction project managers, who did not have any managerial roles, were excluded from the study. The demographic was essential but not the primary eligibility criteria used to include or exclude participants from the study.

3.4.3 Sample Size

One of the significant objectives of sampling the population is to obtain a statistically representative population sample. An appropriate research sample improves the investigator's confidence that the study results could not be affected by random variations in the target population; this explains the importance of the population sample's representativeness in quantitative studies (Majid, 2018). However, according to the eligibility criteria, only managers of construction projects were eligible to participate, and in cases where the managers were unavailable, their immediate supervisors were involved.

The research expected a few participants and, therefore, all managers actively involved in managing the New York City capital projects and other construction projects nationwide involved in the study.

In order to get accurate, statistically significant results, using these criteria, the sample size of the study included 281 project managers, whereby 164 were project managers who work in construction projects nationwide and 117 project managers who were directly involved in managing New York City capital projects. A response rate of 54% translated to 89 respondents from nationwide construction projects and 63 respondents from New York City capital projects.

3.5 Research Process and Data Collection

This refers to the systematic processes involved in conducting the actual study. The research process began by identifying the problem of cost overruns in the construction industry, especially on publicly funded construction processes. A critical review of the existing body of literature (Tezel and Niezel,2013). indicated that numerous methods had been employed to mitigate the problem of cost overruns in construction projects without success over the years. The existing literature has revealed increased adoption of lean principles that have been traditionally used in the manufacturing industry to the construction industry (Bajjou and Chafi,2018). Therefore, this study sought to establish whether the application of lean construction principles reduced cost overruns in the construction industry.

The study identified nine fundamental principles and twenty-five sub-principles that form the basis upon which a survey questionnaire was prepared.

The contacts of the potential respondents were obtained from the New York City capital projects database, and those from other construction projects nationwide were obtained from the respective websites, resulting in a comprehensive database comprising

the name of the project, the name of the manager or supervisor, email addresses and phone numbers of the potential respondents.

Initially, an email was sent to the selected population sample to invite them to participate in the online survey and only those who confirm their participation guided to the survey questionnaire.

The participants who confirmed their participation sent the link directing them to the online survey through the email addresses given. They were required to complete the questionnaire and returned their responses within 21 days. After 14 days, a follow-up by sending reminders to those who did not return their responses was instituted. Those who did not respond by the last day called to determine whether they intend to respond to the questionnaire, only responses received within the stipulated time of 21 days used in the study.

3.6 Research Instruments

The research was carried out by using an online survey. The survey comprised an online questionnaire containing twenty-five close-ended questions. The close-ended questionnaire required respondents to choose the most appropriate options ranked on a five-point Likert scale. The degree of application of each of the twenty -five sub-principles for each specific construction project has been measured on a Likert scale from 1 to 5, where 1 denotes the lowest degree of application (never), while 5 denotes the highest degree of application (always) of the relevant sub-principle. The survey was expected to take 15 to 20 minutes.

Online questionnaires provide a valid, timely, and cost-conscious way of collecting quantitative data from a large, diverse research population (Nigel, Nick, and Amanda, 2009).

A Likert scale provides a helpful tool for measuring the respondents' opinions and attitudes towards the solutions proposed for the problem being addressed by the study (Suvillan and Artino,2013). It eliminates the occurrence of inconsistent responses in the study by allowing respondents to choose only one of the options. The ranking of responses on the Likert scale makes the data analysis process more manageable, as all the responses are represented in numerical values that are uncomplicated to manage using statistical tools (Nigel et al., 2009).

3.7 Item Analysis by Applying Cronbach's Alpha

"Item Analysis explains how well a set of questions (or items) measures one characteristic (or construct) and helps to identify problematic questions." (Griffith, 2015). "A standard method for examining the reliability of individual construct in research is Cronbach's alpha" (George, 2011). "The larger Cronbach's α value ensured the internal consistency among the constructs with an acceptable range >0.70" (Nunnally, 1978 and Hair et al., 2019).

The application of Cronbach's alpha involved analyzing the complete data (the entire survey) and every category of lean principles and sub-principles for reliability.

3.7.1 Cronbach's Alpha

The Cronbach's Alpha analysis by the following equation calculated:

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum \sigma_i^2}{\sigma_i}\right) \quad (1)$$

Where,

σ_i^2 scores' variance for each factor

$\sum \sigma_i^2$ the sum of scores' variance

k is the number of factors

3.8 Data Analyses

In this research, data analyses carried out using the *Random Forest* machine learning in the *R* programming environment and *Partial Least Squares* (PLS) using *Minitab* software.

Research Hypothesis I: Projects that utilize Lean Construction Principles (Including Waste Elimination, Supply and People Involvement, Planning and Scheduling, Standardization, Quality, Transparency, Continuous Improvement, and Customer Focus) tend to have less cost over-runs.

Research Hypothesis II: The application of Lean Construction Principles (LCPs) will significantly reduce cost overruns in New York City construction projects.

Random Forest machine learning algorithm, initially proposed in Breiman L. (2001), was used to identify the most influential factors among all twenty-five sub-principles (independent variables), while *Partial Least Squares* (PLS) was used to build a predictive model to test the two research hypotheses.

"Partial least squares (PLS) regression is a method that reduces the predictors to a smaller set of uncorrelated components and performs least squares regression on these components instead of on the original data. PLS regression is instrumental when the predictors are highly collinear or when there are more predictors than observations, and ordinary least-squares regression either produces coefficients with high standard errors or fails" (Blackburn, 2020).

The *Random Forest* machine learning algorithm is known for its high performance over different data sets, including small sample data sets. That algorithm does not make assumptions about the dependent variable's probability distribution (Breiman, 2001). Also, the Random Forest algorithm's predictive capability is not influenced by the presence of multicollinearity among predictors (Breiman, 2001).

Random Forest algorithm builds a complicated non-linear model that considers all possible interactions between predictors, which cannot be achieved by the *Multiple Regression* procedure and the likes (Breiman, 2001).

Using Random Forest machine learning, the most important predictors(factors) associated with each survey were determined to validate hypothesis I. The entire data set (all questions) randomly selected 70% (Training data) and developed a predictive model for percent over-budget using PLS. Using the equation generated, the percent over-budget with the remaining 30% (Test data) predicted and compared to the actuals.

To validate hypothesis II, the NYC data stratified out. Each Likert score increased by 1 when less than 5, using the PLS model's prediction equation and

predicting the new percent over. Then a Paired T-Test was used to confirm there was an improvement after.

3.8.1 Paired T-Test

The paired t-test analysis by the following equation calculated:

$$t = \frac{m}{S/\sqrt{n}} \quad (2)$$

Where,

m is the mean differences

n is the sample size (i.e., size of d)

S is the standard deviation of *d*

3.9 Ethical Considerations

I. Permission to Conduct Research

The current study will be conducted in strict conformity with ethical standards that apply to quantitative online research and interviews. Written permission to conduct the research will be obtained from the George Washington University. Permission from the project companies to make them aware of the research and the study's nature will also be obtained.

II. Consent to Participate in the Study

The study's purpose will be clarified, including a brief explanation of the procedures and a description of benefits reasonably expected. An instruction that participants will be free to opt-out of the study if they experience any violation will be provided (Creswell, 2013).

III. Risk of Harm, Anonymity, and Confidentiality of the Participants

Per the guidelines of conducting a study, the participants will be asked not to include their names on the form containing their responses. The study will maintain the highest level of confidentiality while handling the participants. There will be no sharing of individual results with the participants. Only aggregated findings will be communicated to those who will express interest in the study (Creswell, 2013). Any potential risk of harm, like resource loss, reputational, physical, or emotional involving the current study will be communicated to the participants and will be considerably minimized.

Chapter 4—Results

4.1 Introduction

The Data analysis chapter presents analyses and discussion of the results of the study. The primary data was collected from a total of 152 responses from construction project managers. The data was cleaned up, coded, and analyzed by applying the *Random Forest* algorithm in the *R* programming environment and *Partial Least Squares* (PLS) with *Minitab* software. This chapter begins with an exploration of the demographic details of the respondents of the study. These include the type of project that the project managers have managed, the region where the project managers have worked in their last construction project. The chapter also presents which predictors and factors associated with each survey question are significant.

Furthermore, the chapter presents the results of the Paired Sample t-Test used to present the differences in terms of cost between the initial project cost and the new project cost after application of Lean Construction Principles (LCPs) in New York City construction projects. The last section of this chapter focuses on the most important factors, the research product, and how engineering managers can use it.

4.2 General Information

4.2.1 Project Type

A total of 152 construction project managers were asked to indicate the type of construction project they have recently managed. Out of 152 responses, 49 (32.2%) of project managers indicated that in their last construction project, they had managed *Commercial & Institutional* projects such as hospitals, clinics, stadiums, sports facilities,

shopping centers, etc. A total of 21 (13.8%) project managers have managed *Residential* projects such as townhouses, houses, condominiums, apartments, cottages, subdivisions, single-unit dwellings, etc. A total of 21 (13.8%) of project managers have also indicated that they were in charge of *Transportations* projects such as the alteration, repair, and construction of roads, streets, alleys, highways, runways, paths, etc. A total of 18 (11.8%) project managers indicated that they had managed *Heavy* projects such as dams, sewer line projects, sewage treatment facilities, dredging projects, flood control projects, water treatment plants, etc.

A total of 16 (10.5%) project managers have indicated that they have been in charge of *Educational* construction projects such as schools and universities. A total of 15(9.9%) project managers have responded that they have worked on *Industrial* projects owned by large industrial corporations such as medicine, power generation, manufacturing, petroleum, mining. Lastly, 12 (7.9%) responses came from project managers in charge of building projects such as small renovations. These results of the survey are summarized in figure 4.1 below:

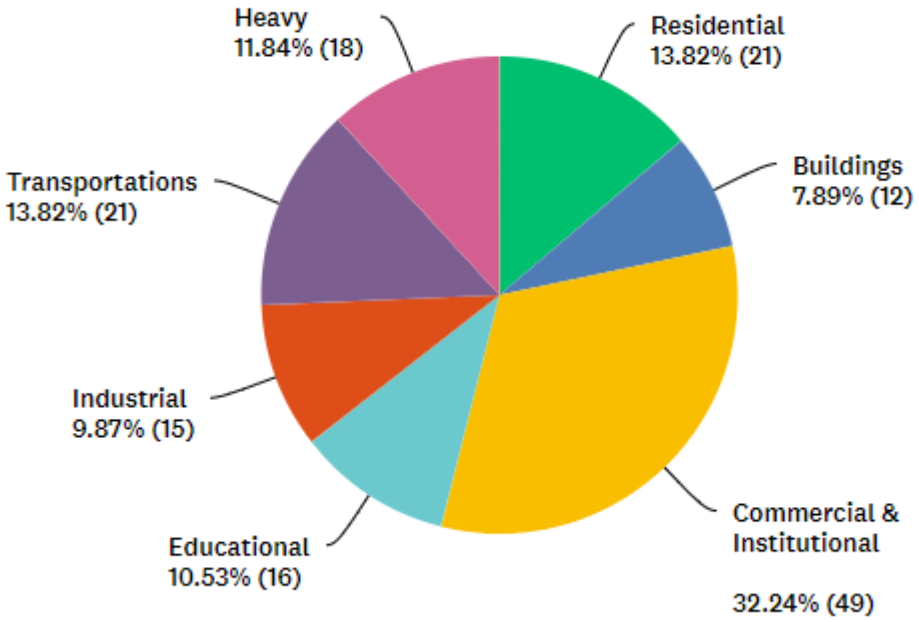


Figure 4.1 Type of Most Recent Projects Managed by Construction Project Managers

4.2.2 The Region of Project Managers' Last Project

The 152 construction project managers were asked to indicate the region where they have managed their last construction projects. 72 (47%) of the project managers' last construction projects were located in the Northeast region, including Maine, New Hampshire, Vermont, Massachusetts, Rhode Island, Connecticut, New York, New Jersey Pennsylvania. 41 (27%) of the project managers' last construction projects were located in the West: Including Montana, Idaho, Wyoming, Colorado, New Mexico, Arizona, Utah, Nevada, California, Oregon, Washington, Alaska, and Hawaii. 26 (17%) of the project managers' last construction projects were located in the Southeast: Including Delaware, Maryland, Virginia, West Virginia, Kentucky, North Carolina, South Carolina, Tennessee, Georgia, Florida, Alabama, Mississippi, Arkansas, Louisiana, Texas, Oklahoma, and Washington, DC. Furthermore, 13(9%) of the project managers' last construction projects

were located in the Midwest: Including Ohio, Michigan, Indiana, Wisconsin, Illinois, Minnesota, Iowa, Missouri, North Dakota, South Dakota, Nebraska, and Kansas.

The survey results in table 4.1 and figure 4.2 below indicate that the study's sample was derived from a diverse population drawn from the construction industry in various regions of the United States.

		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	The Northeast	72	47 %	47%	47%
	The West	41	27%	27%	74%
	The Southeast	26	17%	17%	91%
	The Midwest	13	9%	9%	100%
	Total	152	100%	100%	

Table 4.1 The Region of Construction Project Managers' Last Project

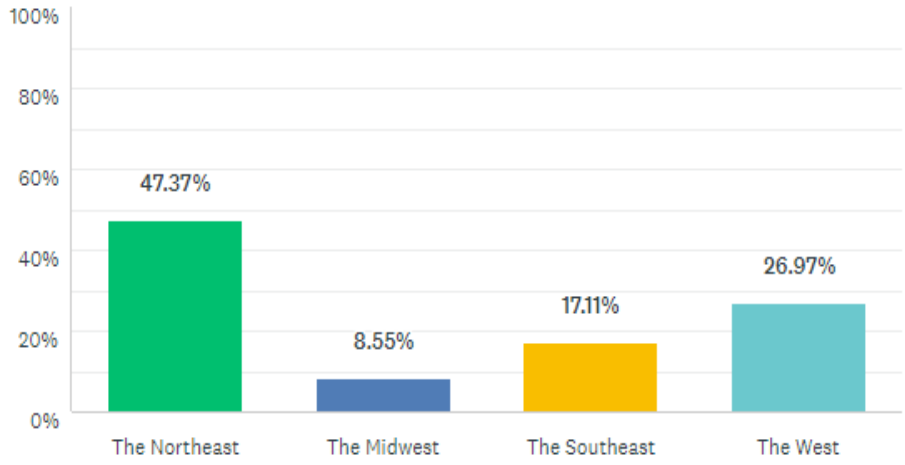


Figure 4.2 The Region of Construction Project Managers' Last Project

4.3 Analysis of Responses Based on the Likert Scale

By choosing one of the responses ranked on the five-point Likert scale, where;

- N5 = Number of respondents who indicated, "Always."
- N4 = Number of respondents who indicated "Often."
- N3 = Number of respondents who indicated "Sometimes."
- N2 = Number of respondents who indicated "Rarely."
- N1 = Number of respondents who indicated "Never."
- A = Highest weight = 5
- N = Total number of respondents = 152

The respondents ascribed the respective importance to the respective lean construction principle or tool. The summary ascribed to each Lean Construction sub-Principles is as follows:

1.1. Supply Chain Management (SCM)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	10	6.6	6.6	6.6
	2.00	31	20.4	20.4	27.0
	3.00	39	25.7	25.7	52.6
	4.00	48	31.6	31.6	84.2
	5.00	24	15.8	15.8	100.0
	Total	152	100.0	100.0	

1.2. Just- In- Time (JIT)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	11	7.2	7.2	7.2
	2.00	37	24.3	24.3	31.6
	3.00	30	19.7	19.7	51.3
	4.00	43	28.3	28.3	79.6
	5.00	31	20.4	20.4	100.0
	Total	152	100.0	100.0	

1.3. Pull System

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	18	11.8	11.8	11.8
	2.00	28	18.4	18.4	30.3
	3.00	49	32.2	32.2	62.5
	4.00	43	28.3	28.3	90.8
	5.00	14	9.2	9.2	100.0
	Total	152	100.0	100.0	

1.4. Suppliers' Involvement

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	10	6.6	6.6	6.6
	2.00	22	14.5	14.5	21.1
	3.00	40	26.3	26.3	47.4
	4.00	57	37.5	37.5	84.9
	5.00	23	15.1	15.1	100.0
	Total	152	100.0	100.0	

2.1. People/Worker Involvement

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	9	5.9	5.9	5.9
	2.00	24	15.8	15.8	21.7
	3.00	41	27.0	27.0	48.7
	4.00	51	33.6	33.6	82.2
	5.00	27	17.8	17.8	100.0
	Total	152	100.0	100.0	

2.2. Teamwork

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	16	10.5	10.5	10.5
	2.00	23	15.1	15.1	25.7
	3.00	36	23.7	23.7	49.3
	4.00	47	30.9	30.9	80.3
	5.00	30	19.7	19.7	100.0
	Total	152	100.0	100.0	

3.1. Metrics

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	13	8.6	8.6	8.6
	2.00	36	23.7	23.7	32.2
	3.00	43	28.3	28.3	60.5
	4.00	35	23.0	23.0	83.6
	5.00	25	16.4	16.4	100.0
	Total	152	100.0	100.0	

3.2 Huddle Meetings

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	4	2.6	2.6	2.6
	2.00	22	14.5	14.5	17.1
	3.00	55	36.2	36.2	53.3
	4.00	49	32.2	32.2	85.5
	5.00	22	14.5	14.5	100.0
	Total	152	100.0	100.0	

4.1. Flexible Resources

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	16	10.5	10.5	10.5
	2.00	22	14.5	14.5	25.0
	3.00	37	24.3	24.3	49.3
	4.00	45	29.6	29.6	78.9
	5.00	32	21.1	21.1	100.0
	Total	152	100.0	100.0	

4.2. Optimize Value

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	9	5.9	5.9	5.9
	2.00	28	18.4	18.4	24.3
	3.00	43	28.3	28.3	52.6
	4.00	48	31.6	31.6	84.2
	5.00	24	15.8	15.8	100.0
	Total	152	100.0	100.0	

4.3. Relationships

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	4	2.6	2.6	2.6
	2.00	24	15.8	15.8	18.4
	3.00	49	32.2	32.2	50.7
	4.00	46	30.3	30.3	80.9
	5.00	29	19.1	19.1	100.0
	Total	152	100.0	100.0	

5.1. Optimize Product System

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	19	12.5	12.5	12.5
	2.00	34	22.4	22.4	34.9
	3.00	38	25.0	25.0	59.9
	4.00	44	28.9	28.9	88.8
	5.00	17	11.2	11.2	100.0
	Total	152	100.0	100.0	

5.2. TIMWOOD

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	40	26.3	26.3	26.3
	2.00	48	31.6	31.6	57.9
	3.00	37	24.3	24.3	82.2
	4.00	18	11.8	11.8	94.1
	5.00	9	5.9	5.9	100.0
	Total	152	100.0	100.0	

5.3. Waste Disposal Management

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	8	5.3	5.3	5.3
	2.00	24	15.8	15.8	21.1
	3.00	61	40.1	40.1	61.2
	4.00	41	27.0	27.0	88.2
	5.00	18	11.8	11.8	100.0
	Total	152	100.0	100.0	

5.4 Value Stream Mapping (VSM)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	32	21.1	21.1	21.1
	2.00	47	30.9	30.9	52.0
	3.00	39	25.7	25.7	77.6
	4.00	29	19.1	19.1	96.7
	5.00	5	3.3	3.3	100.0
	Total	152	100.0	100.0	

6.1 Optimize Work Content

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	8	5.3	5.3	5.3
	2.00	28	18.4	18.4	23.7
	3.00	53	34.9	34.9	58.6
	4.00	47	30.9	30.9	89.5
	5.00	16	10.5	10.5	100.0
	Total	152	100.0	100.0	

6.2. Takt Time

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	41	27.0	27.0	27.0
	2.00	52	34.2	34.2	61.2
	3.00	34	22.4	22.4	83.6
	4.00	15	9.9	9.9	93.4
	5.00	10	6.6	6.6	100.0
	Total	152	100.0	100.0	

6.3. Work Sequence

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	1	.7	.7	.7
	2.00	25	16.4	16.4	17.1
	3.00	45	29.6	29.6	46.7
	4.00	57	37.5	37.5	84.2
	5.00	24	15.8	15.8	100.0
	Total	152	100.0	100.0	

7.1 Percent Plan Complete Indicator (PPC)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	16	10.5	10.5	10.5
	2.00	33	21.7	21.7	32.2
	3.00	50	32.9	32.9	65.1
	4.00	39	25.7	25.7	90.8
	5.00	14	9.2	9.2	100.0
	Total	152	100.0	100.0	

7.2 Collaborative Planning

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	10	6.6	6.6	6.6
	2.00	32	21.1	21.1	27.6
	3.00	41	27.0	27.0	54.6
	4.00	50	32.9	32.9	87.5
	5.00	19	12.5	12.5	100.0
	Total	152	100.0	100.0	

8.1 Total Quality Management (TQM)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	18	11.8	11.8	11.8
	2.00	23	15.1	15.1	27.0
	3.00	52	34.2	34.2	61.2
	4.00	46	30.3	30.3	91.4
	5.00	13	8.6	8.6	100.0
	Total	152	100.0	100.0	

8.2 Error-Proofing (Poka-Yoke)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	47	30.9	30.9	30.9
	2.00	58	38.2	38.2	69.1
	3.00	26	17.1	17.1	86.2
	4.00	16	10.5	10.5	96.7
	5.00	5	3.3	3.3	100.0
	Total	152	100.0	100.0	

8.3 The Five Whys

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	6	3.9	3.9	3.9
	2.00	24	15.8	15.8	19.7
	3.00	49	32.2	32.2	52.0
	4.00	56	36.8	36.8	88.8
	5.00	17	11.2	11.2	100.0
	Total	152	100.0	100.0	

9.1 Work Organization

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	15	9.9	9.9	9.9
	2.00	43	28.3	28.3	38.2
	3.00	52	34.2	34.2	72.4
	4.00	31	20.4	20.4	92.8
	5.00	11	7.2	7.2	100.0
	Total	152	100.0	100.0	

9.2 Building Information Modeling (BIM)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	40	26.3	26.3	26.3
	2.00	57	37.5	37.5	63.8
	3.00	38	25.0	25.0	88.8
	4.00	11	7.2	7.2	96.1
	5.00	6	3.9	3.9	100.0
	Total	152	100.0	100.0	

4.4 Summery of the Frequencies of the Responses

This section summarizes the frequencies of the online survey responses to determine the implementation of lean construction sub- principles in the United States' construction industry. A summary of the responses collected is provided in the table below:

Sub-Principles	Never	Rarely	Sometimes	Often	Always
1.1. Supply Chain Management (SCM)	10	31	39	48	24
1.2. Just- In- Time (JIT)	11	37	30	43	31
1.3. Pull System	18	28	49	43	14
1.4. Supplier Involvement	10	22	40	57	23
2.1. People/Worker Involvement	9	24	41	51	27
2.2. Teamwork	16	23	36	47	30
3.1. Metrics	13	36	43	35	25
3.2 Huddle Meetings	4	22	55	49	22
4.1. Flexible Resources	16	22	37	45	32
4.2. Optimize Value	9	28	43	48	24
4.3. Relationships	4	24	49	46	29
5.1. Optimize Product System	19	34	38	44	17
5.2. TIMWOOD	40	48	37	18	9
5.3. Waste Disposal Management	8	24	61	41	18
5.4 Value Stream Mapping (VSM)	32	47	39	29	5
6.1 Optimize Work Content	8	28	53	47	16
6.2. Takt Time	41	52	34	15	10
6.3. Work Sequence	1	25	45	57	24
7.1 Percent Plan Complete Indicator (PPC)	16	33	50	39	14
7.2 Collaborative Planning	10	32	41	50	19
8.1 Total Quality Management (TQM)	18	23	52	46	13
8.2 Error-Proofing (Poka-Yoke)	47	58	26	16	5
8.3 The Five Whys	6	24	49	56	17
9.1 Work Organization	15	43	52	31	11
9.2 Building Information Modeling (BIM)	40	57	38	11	6
Total	11%	22%	28%	27%	12%

Table 4.2 Summary of the Frequencies of the Responses

The summary of the frequencies for lean construction principles, as indicated in the table above, shows that 11% of project managers have never used the Lean Construction Principles (LCPs) listed in this study. 22% of the responses showed that project managers rarely used these principles. 28% of responses affirm that project managers were somewhat familiar with the sub-principles. Also, 27% of responses indicate that the project managers acknowledged the lean construction sub-principles. Furthermore, 12 % of the responses suggest that project managers well recognized lean construction sub-principles.

Cumulatively, many responses showing that project managers were familiarized with lean construction principles. Although it is a relatively new construction concept, project managers increasingly adopt Lean Construction Principles (LCPs). Nevertheless, many responses also indicated a lack of familiarity or insufficiency in implementing these principles.

4.5 Item Analysis

"Item Analysis explains how well a set of questions (or items) measures one characteristic (or construct) and helps to identify problematic questions." (Griffith, 2015). "A standard method for examining the reliability of individual construct in research is Cronbach's alpha" (George, 2011). "The larger Cronbach's α value ensured the internal consistency among the constructs with an acceptable range >0.70" (Nunnally, 1978 and Hair et al., 2019). The application of Cronbach's alpha involved analyzing the complete data (the entire survey) and every category of lean principles and sub-principles for reliability. Reliability analysis results are shown below:

Principle	Sub- principles	Cronbach's Alpha	Overall Cronbach's Alpha
1. Supply	1.1. Supply Chain Management (SCM)	.961	.962
	1.2. Just- In- Time (JIT)	.959	
	1.3. Pull System	.960	
	1.4. Supplier Involvement	.961	
2. People Involvement	2.1. Worker's Involvement	.959	
	2.2. Teamwork	.960	
3.Continuous Improvement	3.1. Metrics	.960	
	3.2 Huddle Meetings	.960	
4. Customer Focus	4.1. Flexible Resources	.959	
	4.2. Optimize Value	.959	
	4.3. Relationships	.960	
5. Waste Elimination	5.1. Optimize Product System	.959	
	5.2. TIMWOOD	.960	
	5.3. Waste Disposal Management	.960	
	5.4 Value Stream Mapping (VSM)	.962	
6. Standardization	6.1 Optimize Work Content	.960	
	6.2. Takt Time	.961	
	6.3. Work Sequence	.961	
7. Planning and Scheduling	7.1 Percent Plan Complete Indicator (PPC)	.960	
	7.2 Collaborative Planning	.961	
8. Quality	8.1 Total Quality Management (TQM)	.960	
	8.2 Error-Proofing (Poka-Yoke)	.962	
	8.3 The Five Whys	.960	
9. Transparency	9.1 Work Organization	.960	
	9.2 Building Information Modeling (BIM)	.962	

Table 4.3 Cronbach's Alpha Values

Hair et al. (2019) defined normality as the data distribution's shape for a metric variable and its correspondence to the normal distribution, the benchmark for statistical methods. The data normality assumptions were met.

4.6 Statistical Methods

In this research, data analyses carried out using the *Random Forest* machine learning in the *R* programming environment and *Partial Least Squares* (PLS) using *Minitab* software.

Research Hypothesis I: Projects that utilize Lean Construction Principles (Including Waste Elimination, Supply and People Involvement, Planning and Scheduling, Standardization, Quality, Transparency, Continuous Improvement, and Customer Focus) tend to have less cost over-runs.

Research Hypothesis II: The application of Lean Construction Principles (LCPs) will significantly reduce cost overruns in New York City construction projects.

Random Forest machine learning algorithm, initially proposed in Breiman (2001), was used to identify the most influential factors among all twenty-five sub-principles (independent variables), while *Partial Least Squares* (PLS) was used to build a predictive model to test the two research hypotheses. In this research, the PLS technique in the *Minitab* softwar*e* was used. Using Random Forest machine learning, the most important predictors(factors) associated with each survey were determined to validate hypothesis I. The entire data set (all questions) randomly selected 70% (Training data) and developed a predictive model for percent over-budget. Using the equation generated, the percent over-budget with the remaining 30% (Test data) predicted and compared to the actuals.

To validate research hypothesis II, the NYC data was stratified out. Each Likert score was increased by 1 when less than 5, using the prediction equation and predicting the new percent over. Then traditional *Paired t-test* was used to confirm there was an

improvement after. Some graphical output was provided as additional evidence in favor of the above two research hypotheses.

4.7 Building the Random Forest Model

Random Forest machine learning algorithm, initially proposed in Breiman L. (2001), was used to identify the most important factors among all twenty-five sub-principles (independent variables). One of the main concerns in applying machine learning algorithms is to avoid the so-called O*verfitting ((*Breiman, 2001). When Overfitting occurs, the model under consideration will perform very well on predictions from the training set used to build the model but will have poor performance when used to obtain new data predictions (Breiman, 2001). A common technique used in tuning the machine models to avoid Overfitting is *cross-validation* (Breiman, 2001).

The original data set is partitioned into two non-overlapping data sets in the cross-validation approach: a training set used to fit the model, and a test (validation) set used to assess the model performance. That process can be repeated using different random partitions of the original data set, called *repeated cross-validation* (Breiman, 2001).

The *Random Forest* predictive model's main performance parameter is the *number* of the *randomly* selected predictors (*mtry*) used for building the individual decision trees included in the random forest (Breiman, 2001). The total *number of trees* (*ntree)* used in the algorithm should be selected large enough to achieve the algorithm's robust good performance over a broad range of *mtry* values. For finding the optimal value for *mtry*, the *caret* package from *R* was used. A 10-fold ten times repeated cross-

validation with *ntree* = 1000 was performed for *mtry* =1, 2, and 15, with the result shown in the figure below. Each value of *mtry* 100 different random forest models was fit and tested for their new data performance.

The Random Forest predictive model's performance parameter with quantitative dependent variable used to find the optimal value for *mtry* is *RMSE* (Root Mean Squared Error). Figure 4.3 shows that the minimum value for the *RMSE* is achieved for *mtry* = 3. The best value *mtry* = 3 and *ntree* = 1000 were used to fit the suitable model from the whole data set using a Random Forest package in *R*.

The model is a complex digital object represented neither by an algebraic equation nor by a plot. The model obtained was saved within the *R* workspace as an *R* object named *Random Forest Model*, used for the subsequent predictions.

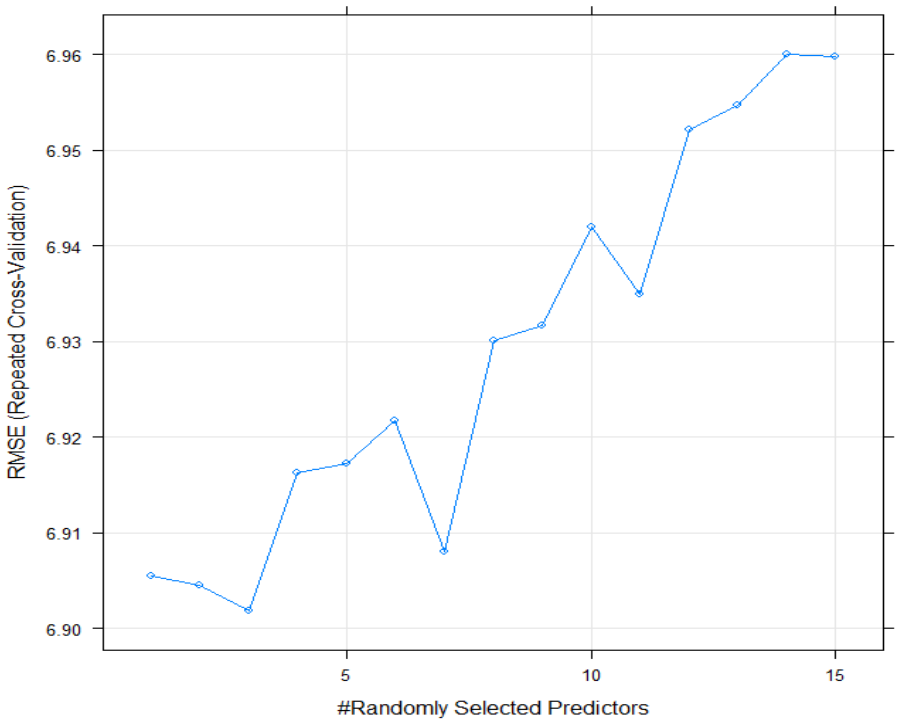
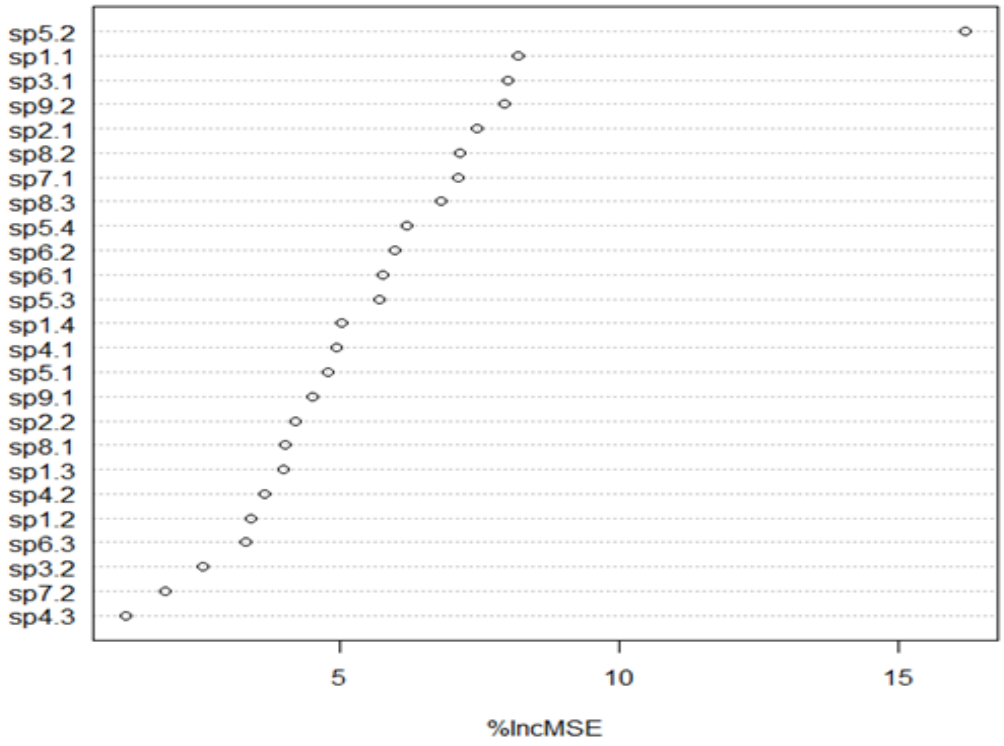


Figure 4.3 Repeated Cross-Validation Results

A notable Variable Importance metric *%IncMSE to* determine the *Relative Variable Importance* of all twenty-five predictors included in the best model was calculated. That metric shows the percent increase in the *MSE* (Mean Square Error) predicting the dependent variable due to removing a specific predictor from the model. In other words, metric *%IncMSE* shows by how much (in percent) the accuracy of the prediction was affected by each of the predictors included in the model. A Relative Variable Importance plot was created and shown in figure 4.4 below. To obtain a neat graphical output, the individual predictors were given short symbolic names sp1.1, sp1.2, …, sp9.2, which correspond to the long names of the twenty-five sub-principles from the original data set.



Random Forest Model

Figure 4.4 Relative Variable Importance Plot

The above results indicated that out of the twenty-five sub- principles of lean construction, TIMWOOD *(5.2)* emerged as the leading sub-principle of perceived importance, followed by Supply Chain Management (1.1), Metrics (3.1), Building Information Modeling (9.2), Workers' Involvement (2.1), Poka-Yoke (8.2), Percent Plan Complete (7.1), The Five Whys (8.3), Value Stream Mapping (5.4) and Takt Time (6.2). The complete ranking of sub-principles based on their importance is shown in the table below:

Principles	*Sub-principles*	*Rank*
1. Supply	1.1 Supply Chain Management (SCM)	2
	1.2 Just-In-Time (JIT)	21
	1.3 Pull System	19
	1.4 Suppliers' Involvement	13
2. People Involvement	2.1 Workers' Involvement	5
	2.2 Teamwork	17
3. Continuous Improvement	3.1 Metrics	3
	3.2 Huddle Meetings	23
4. Customer Focus	4.1 Flexible Resources	14
	4.2 Optimize Value	20
	4.3 Relationships	25
5. Waste Elimination	5.1 Optimize Product System	15
	5.2 TIMWOOD	1
	5.3 Waste Disposal Management	12
	5.4 Value Stream Mapping (VSM)	9
6. Standardization	6.1 Optimize Work Content	11
	6.2 Takt Time	10
	6.3 Work Sequence	22
7. Planning and Scheduling	7.1 Percent Plan Complete (PPC)	7
	7.2 Collaborative Planning	24
8. Quality	8.1 Total Quality Management (TQM)	18
	8.2 Error-Proofing (Poka-Yoke)	6
	8.3 The Five Whys	8
9. Transparency	9.1 Work Organization	16
	9.2 Building Information Modeling (BIM)	4

Table 4.4 Relative Variable Importance Ranking

It is known (Breiman, 2001) that the Random Forest model's predictive capability is not influenced even in the presence of perfect multicollinearity among the predictors (correlation coefficient equal to 1 in absolute value); however, the presence of the perfectly correlated predictors will influence the *Relative Variable Importance* of those predictors (Breiman, 2001). To check for possible multicollinearity among the predictors included in the model, a correlation matrix was created and visualized in the figure below:

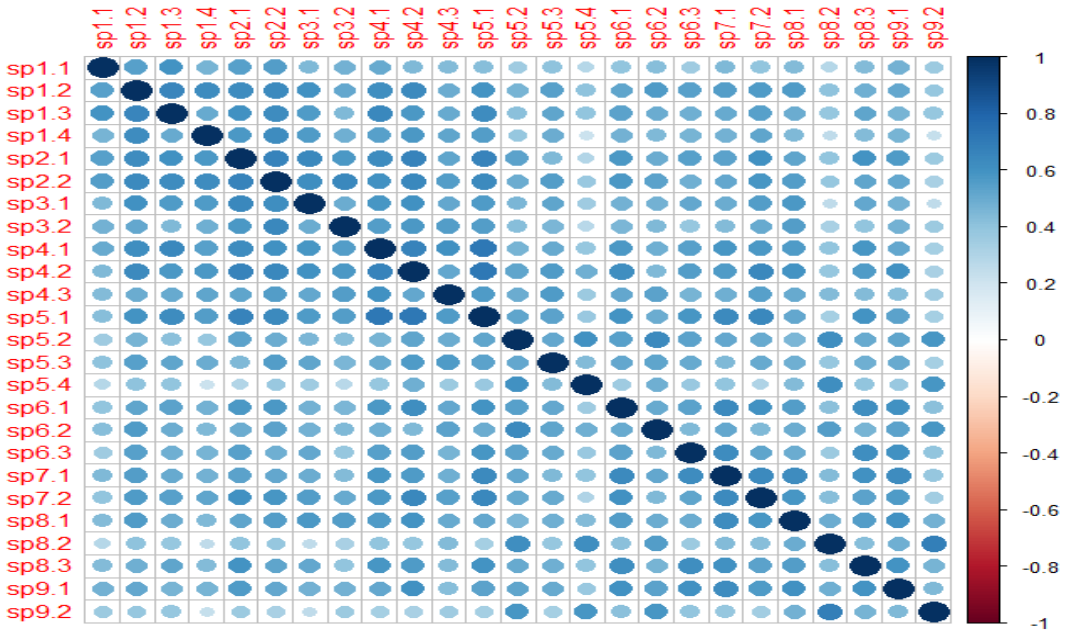


Figure 4.5 Correlation Matrix for Predictors

The above plot reveals a moderate positive correlation between predictors indicating that no multicollinearity among the predictors included in the model, and the Relative Variable Importance Plot in figure 4.5 above is legitimate.

4.8 Partial Least Squares (PLS)

"Partial least squares (PLS) regression is a method that reduces the predictors to a smaller set of uncorrelated components and performs least squares regression on these components instead of on the original data. PLS regression is instrumental when the predictors are highly collinear, or there are more predictors than observations; ordinary least-squares regression either produces coefficients with high standard errors or fails" (Blackburn, 2020). PLS describes the relationship between a set of predictors and one or more continuous responses.

"PLS is useful when there is multicollinearity between predictors. Also, as is often the case with survey data, several predictors will measure similar things, resulting in multicollinearity" (Blackburn, 2020).

PLS method has proved (Mevik and Wehrens,2000) to successfully reduce the number of the new (transformed) variables in the PLS predictive model (called components) in cases where the original factors are correlated with each other and or multicollinearity between factors is present. "PLS helps traditional techniques work or when other machine learning methods do not provide a ready equation when needing the equation to predict" (Blackburn, 2020).

4.9 Partial Least Squares (PLS) Model

A model selection table was used to determine the optimal number of components required to explain the data. The *X*-variance, R^2, and predicted R^2 values in the model selection table were examined. The presence of outliers and leverage points in the data

set was assessed using residual plots (e.g., residual vs. leverage plot) and response plot. This was an essential step in the analysis to confirm high prediction accuracy in the model. Response plot was also used to assess whether a nonlinear pattern is present in the data, indicating misfit in model or prediction inaccuracies. Hence, the model selection was based on *p*-value, *X*-variance, R^2, and predicted R^2 values in the model and the diagnostic plots. A complete model with all the twenty-five sub-principles as independent variables was first assessed.

Finally, the model was validated by using a test data set. The model for hypothesis I was trained using a training data set, 70% of the original data set. Furthermore, the 30% test data set was used to validate the model. For hypothesis II, an augmented data set, through increasing each Likert score by 1 when less than 5 for each sub-principle, was used as the test data set. Then a paired t-test was used to assess whether an improvement in the score for sub-principles would result in an improvement in cost over-run.

4.10 Considerations Regarding Research Hypothesis I

Research Hypothesis I: Projects that utilize Lean Construction Principles (Including Waste Elimination, Supply and People Involvement, Planning and Scheduling, Standardization, Quality, Transparency, Continuous Improvement, and Customer Focus) tend to have less cost over-runs.

Using Random Forest machine learning, the most important predictors(factors) associated with each survey were determined to validate hypothesis I. The entire data set (all questions) randomly selected 70% (Training data) and developed a predictive model

for percent over-budget (using PLS). Using the equation generated, the percent over-budget with the remaining 30% (Test data) predicted and compared to the actuals.

4.10.1 The Model

In the preliminary analysis using the entire model, an outlier was detected (Obs 101), and it was removed from the data set for further analysis.

The number of components selected was *two* using leave-one-out cross-validation.

Cross-validation	Leave-one-out
Components to evaluate	Set
Number of components evaluated	10
Number of components selected	2

Source	DF	SS	MS	F	P
Regression	2	2454.60	1227.30	40.49	0.000
Residual Error	100	3030.93	30.31		
Total	102	5485.53			

Table 4.5 Analysis of Variance Table for Over or Under the Budget Percentage

The model explains the variability in the data adequately ($p<0.001$) at a 0.05 level of significance.

Components	X Variance	Error	R-Sq	PRESS	R-Sq (pred)
1	0.544495	4072.50	0.257592	4292.61	0.217467
2	0.607041	3030.93	0.447468	3851.90	0.297807
3		2545.73	0.535920	3958.69	0.278339
4		2436.25	0.555878	4236.32	0.227728
5		2402.16	0.562091	4458.04	0.187309
6		2395.40	0.563324	4489.08	0.181650
7		2393.41	0.563687	4472.70	0.184637
8		2392.71	0.563815	4462.23	0.186544
9		2392.50	0.563853	4452.81	0.188263
10		2392.46	0.563859	4451.87	0.188434

Table 4.6 Model Selection and Validation for Over or Under Budget Percentage

The model explains 61% of the variance in the terms that the model explains. Hence the components represent the original set of terms adequately. The R^2 value of 45% corresponds to the variation in the response explained by the selected model. The relatively low predicted R^2 indicates that the predictive ability of the model might be

low. The considerable difference between R^2 and predicted R^2 may hint at an overfit in the model.

Variable	Coefficients
Intercept	15.7758
Supply Chain Management (SCM)	-0.3473
Just- In- Time (JIT)	-0.259
Pull System	-0.0056
Supplier Involvement	-0.3323
People/Worker Involvement	0.5666
Teamwork	0.037
Metrics	0.6407
Huddle Meetings	0.0846
Flexible Resources	0.3177
Optimize Value	0.3403
Relationships	-0.0141
Optimize Product System	0.1877
TIMWOOD	-1.6203
Waste Disposal Management	-0.2335
Value Stream Mapping (VSM)	-0.2441
Optimize Work Content	0.4252
Takt Time	-1.4455
Work Sequence	-0.7495
Percent Plan Complete Indicator (PPC)	-0.5238
Collaborative Planning	0.0054
Total Quality Management (TQM)	0.1717
Error-Proofing (Poka-Yoke)	-0.9026
The Five Whys)	0.13
Work Organization	0.133
Building Information Modeling (BIM)	-1.1953

Table 4.7 Model Coefficients

By interpreting the regression model's coefficients while keeping other variables constant, *one* unit increase in the following variables' Likert scale response will have more than 0.5% improvement (reduction) in cost overrun.

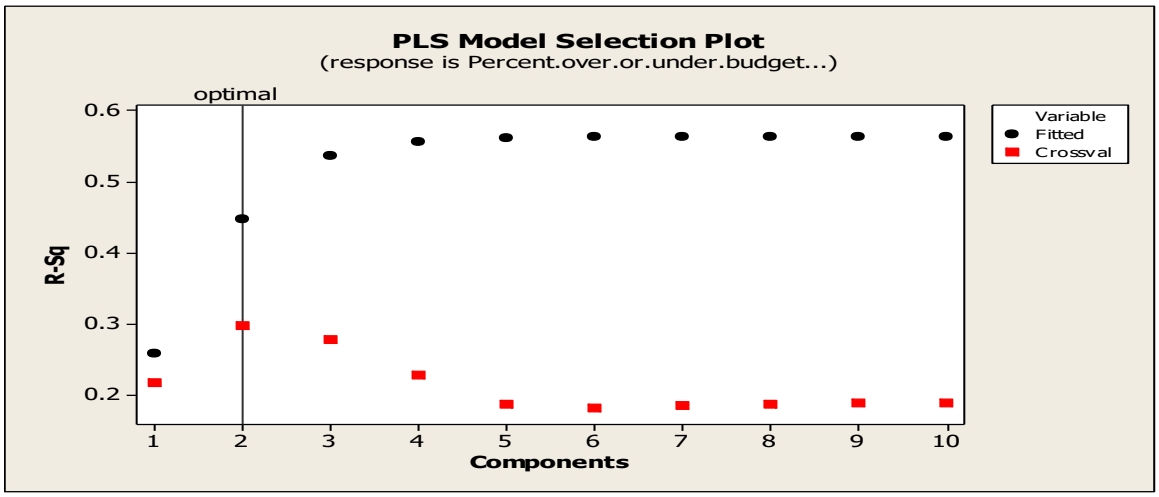


Figure 4.6 Model Selection Plot

The predicted R^2 is maximum at the *two* components.

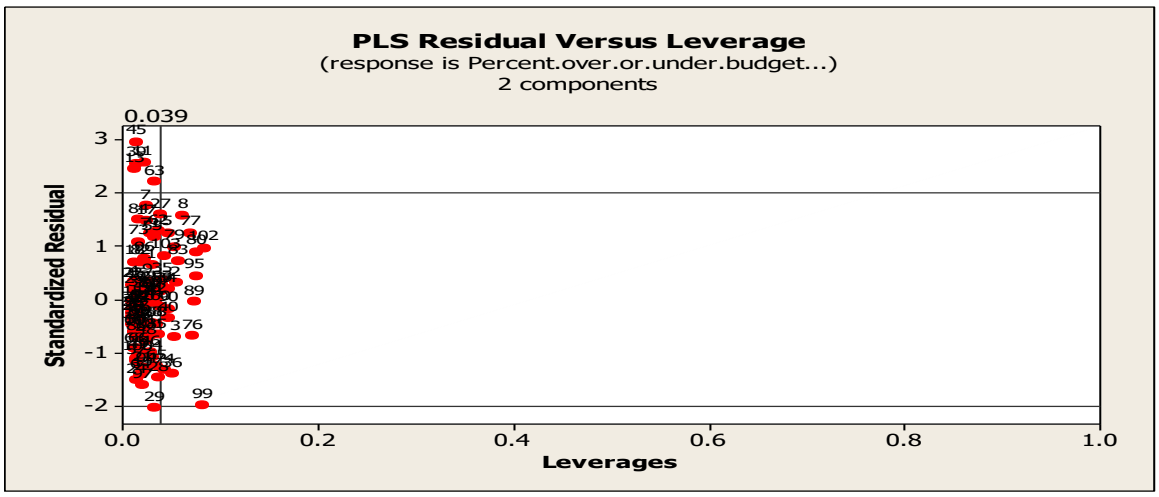


Figure 4.7 Residual versus Leverage Plot

There are no more outliers or leverage points with a drastic difference from the other points after removing the 101st row of data.

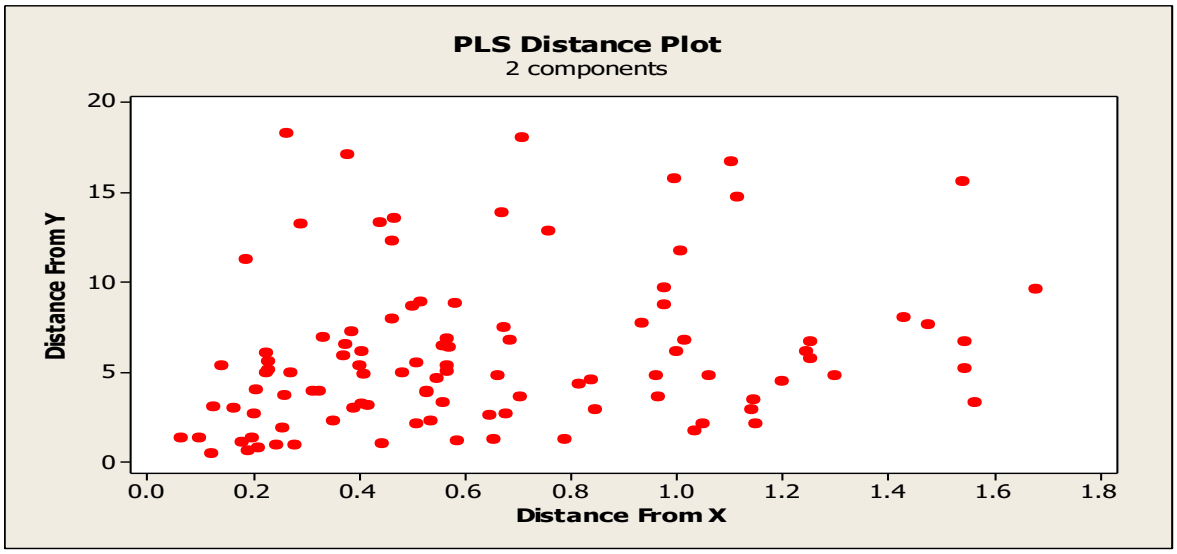


Figure 4.8 Distance Plot

The distance plot does not show any substantially significant leverage points or outliers.

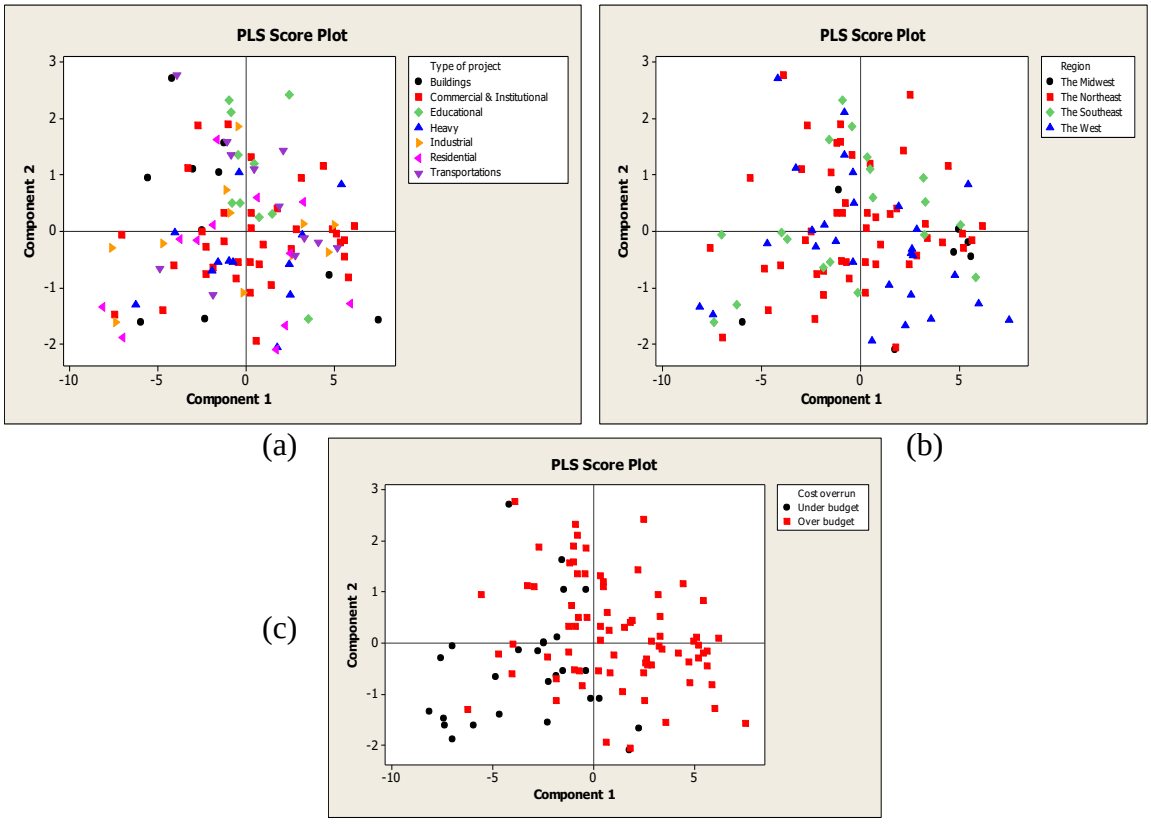


Figure 4.9 Score Plot Categorized by (a) Type of Project (b) Region and (c) Over or Under Budget of Cost Overrun

According to the two external factors, the data does not show any clustering, *project type*, and *region*. However, an apparent clustering can be seen according to the internal factor *over* or *under budget* of cost overrun.

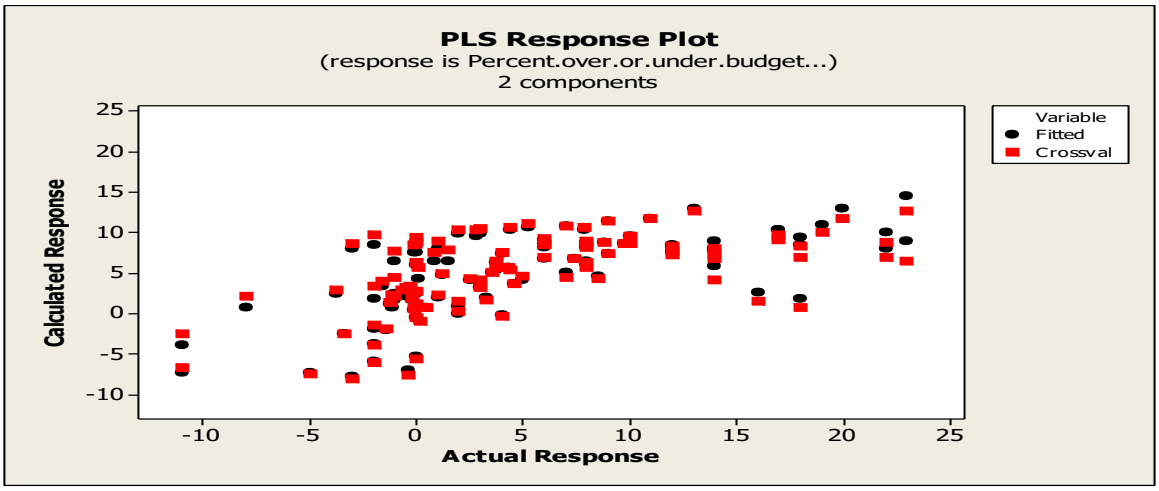


Figure 4.10 Response Plot

There is not much difference between the observed values and the predicted values, indicating a good fit, and there are no more leverage points or outliers in the data.

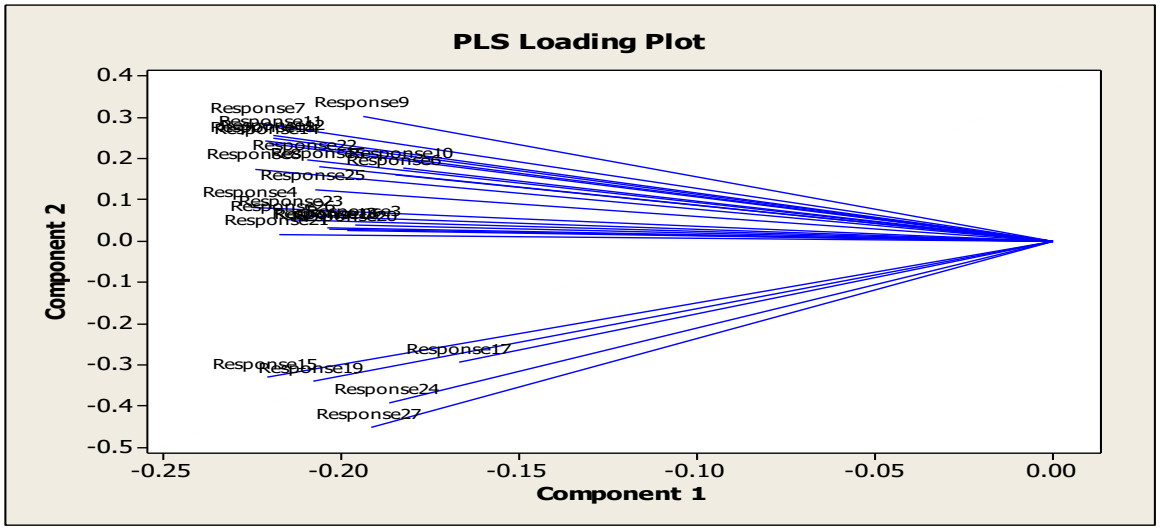


Figure 4.11 Loading Plot

The small angles between the lines of each cluster show the high correlation between variables. The lines are almost the same length, indicating the predictors are

equally important. However, there are some lines seemingly shorter than others. There are possibly two clusters of sub principals with regards to the correlation with each other. The first cluster of variables has positive values, and the second cluster has negative values for the second component, respectively.

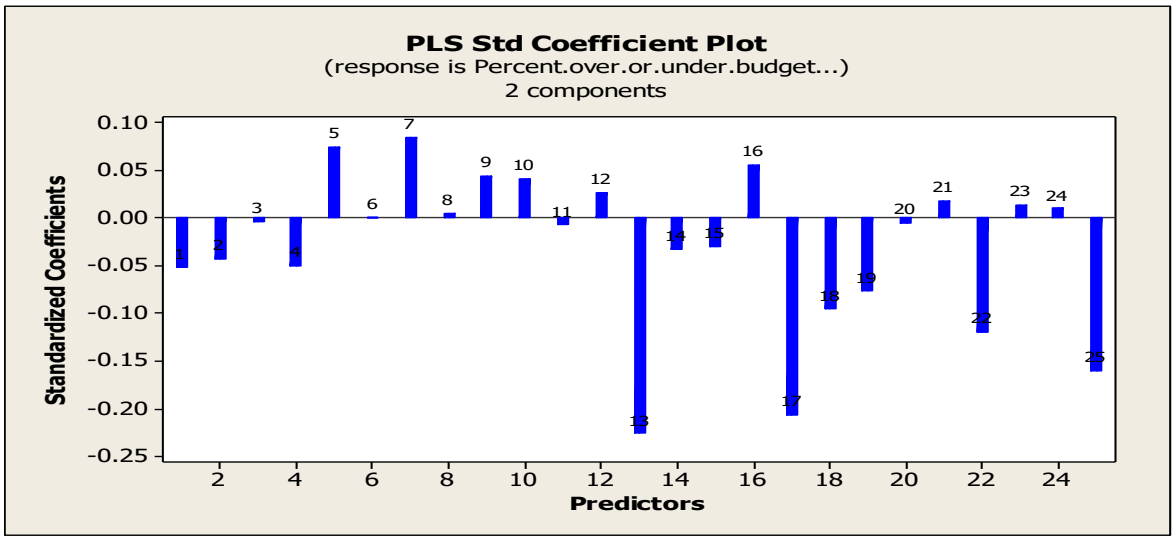


Figure 4.12 Coefficient Plot

While most of the sub-principles are negatively associated with the cost overrun (making the cost overrun improve), some sub- principles are positively associated.

0.35% reduction by 1 unit increase in Supply Chain Management
0.26% reduction by 1 unit increase in Just- In- Time (JIT)
0.01% reduction by 1 unit increase in Pull System
0.33% reduction by 1 unit increase in Suppliers' Involvement
0.01% reduction by 1 unit increase in Relationships
1.62% reduction by 1 unit increase in TIMWOOD
0.23% reduction by 1 unit increase in Waste Disposal Management
0.24% reduction by 1 unit increase in Value Stream Mapping (VSM)
1.45% reduction by 1 unit increase in Takt Time
0.75% reduction by 1 unit increase in Work Sequence
0.52% reduction by 1 unit increase in Percent Plan Complete Indicator
0.90% reduction by 1 unit increase in Error-Proofing (Poka-Yoke)
1.20% reduction by 1 unit increase in Building Information Modeling (BIM)

Table 4.8 Sub-Principles which are Negatively Associated with the Cost Over-run

4.10.2 Final PLS Model

The final session window output for the final PLS model is shown below:

Method

Cross-validation	Leave-one-out
Components to evaluate	Set
Number of components evaluated	10
Number of components selected	2

Analysis of Variance for Percent Over or Under Budget

Source	DF	SS	MS	F	P
Regression	2	2454.60	1227.30	40.49	0.000
Residual Error	100	3030.93	30.31		
Total	102	5485.53			

Model Selection and Validation for Percent Over and Under Budget

Components	X Variance	Error	R-Sq	PRESS	R-Sq (pred)
1	0.544495	4072.50	0.257592	4292.61	0.217467
2	0.607041	3030.93	0.447468	3851.90	0.297807
3		2545.73	0.535920	3958.69	0.278339
4		2436.25	0.555878	4236.32	0.227728
5		2402.16	0.562091	4458.04	0.187309
6		2395.40	0.563324	4489.08	0.181650
7		2393.41	0.563687	4472.70	0.184637
8		2392.71	0.563815	4462.23	0.186544
9		2392.50	0.563853	4452.81	0.188263
10		2392.46	0.563859	4451.87	0.188434

Table 4.9 Final PLS Model

The model explained 60.7% of the variance in the terms that the model explained. The R^2 value of 44.7% corresponds to the variation in the response explained by the selected model.

4.10.3 The Model Validation

I. Two-Sample T-Test

Predictions for the cost overrun of the 30% test data set were calculated in excel using the final PLS equation derived from hypothesis I. A two-sample T-test was performed to compare the actual cost overrun and the predicted cost overrun of the 30%

test data set. There was insufficient evidence to reject the null hypothesis and conclude the means are unequal at 0.05 level of significance (p=0.091). The results are shown in the table below:

Two-sample T for Actual Cost overrun 30% vs. Predictions 30%				
	N	Mean	StDev	SE Mean
Actual Cost overrun 30%	48	7.06	8.53	1.2
Predictions 30%	48	4.65	4.68	0.68
Difference = mu (Actual Cost overrun 30%) - mu (Predictions 30%) Estimate for difference: 2.41 95% CI for difference: (-0.39, 5.21) T-Test of difference = 0 (vs not =): T-Value = 1.71 P-Value = 0.091 DF = 72				

Table 4.10 Two-Sample T-Test and CI: Actual Cost overrun 30%, Predictions 30%

II. Equivalence Test

Equivalence test will be used with paired data to evaluate whether a test population's mean equivalent to a reference population's mean when there are paired (dependent) observations (Wachs, 2015). To further support equivalency, a two-sample equivalency test was run **to** test whether the actual cost overrun of the 30% test data and predictions are equivalent to within ±5. Equivalence at 0.05 level of significance (p=0.0.035) can be claimed. The results are shown in the table below:

Two-Sample Equivalence Test

Descriptive Statistics

Variable	N	Mean	StDev	SE Mean
Actual Cost overrun 30%	48	7.0583	8.5292	1.2311
Predictions 30%	48	4.6507	4.6769	0.67506

Difference: Mean (Test) - Mean (Reference)

Difference	SE	95% CI for Equivalence	Equivalence Interval
-2.4100	1.4043	(-4.75003,0)	(-5, 5)

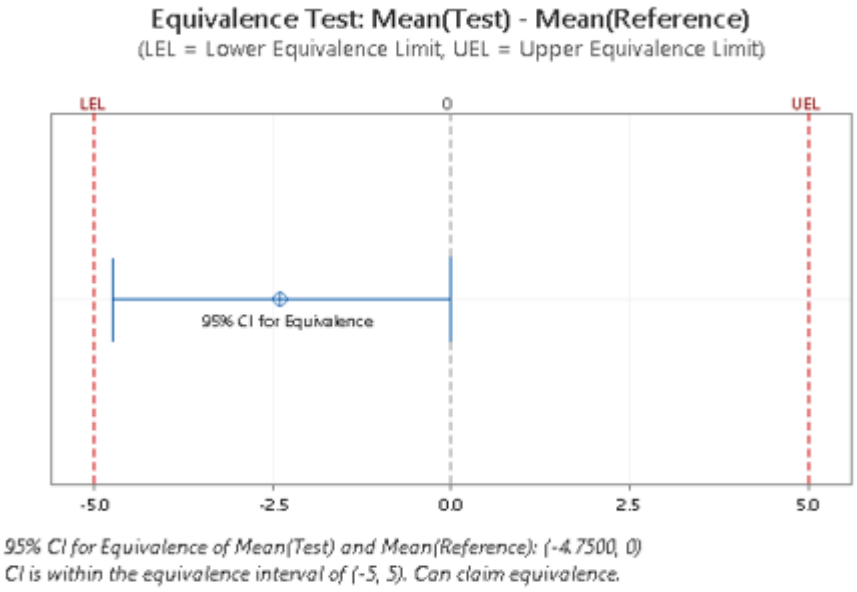
CI is within the equivalence interval. Can claim equivalence.

Test

Null hypothesis:	Difference ≤ -5 or Difference ≥ 5
Alternative hypothesis:	-5 < Difference < 5
α level:	0.05

Null Hypothesis	DF	T-Value	P-Value
Difference ≤ -5	72	1.8443	0.035
Difference ≥ 5	72	-5.2765	0.000

The greater of the two P-Values is 0.035. Can claim equivalence.



95% CI for Equivalence of Mean(Test) and Mean(Reference): (-4.7500, 0)
CI is within the equivalence interval of (-5, 5). Can claim equivalence.

Table 4.11 Two-Sample Equivalence Test for Comparing Means of Actual Cost overrun 30% and Predictions 30%

4.11 Considerations Regarding Research Hypothesis II

Research Hypothesis II: The application of Lean Construction Principles (LCPs) will significantly reduce cost overruns in New York City construction projects.

To validate the research hypothesis II, the NYC data stratified out. Each Likert score increased by 1 when less than 5, using the equation generated by hypothesis I and predicting the new percent over. Then a Paired T-Test was used to confirm there was an improvement after.

Paired t-test was carried out to test the null hypothesis, 'there has been no improvement in cost over-run with an improvement (improvement is a reduction) in the Likert scale.' Furthermore, we have enough evidence to reject the null hypothesis ($p<0.001$) at a 0.05 level of significance and conclude a decrease in cost over-run with an improvement in the Likert scale. Paired t-test for NYC projects cost overrun - Predicted cost overrun for improved Likert scores are shown in the table below:

	N	Mean	StDev	SE Mean
NYC projects cost overrun	63	7.24	7.96	1.00
Predicted cost overrun for improved Likert scores	63	0.62	4.03	0.51
Difference	63	6.621	6.927	0.873
95% lower bound for mean difference: 5.164 T-Test of mean difference = 0 (vs > 0): T-Value = 7.59 P-Value = 0.000				

Table 4.12 Paired T-Test and CI for the Observed Cost Over-run in NYC Projects and Predicted Cost Over-runs after Improvement in the Likert Scale Questions (Using the Augmented Data Set)

	N	Mean	StDev	SE Mean
Cost Overrun in NYC	63	7.24	7.96	1.00
Cost Overrun in USA	152	5.61	7.76	0.63

Table 4.13 Two-Sample T-Test for Comparing Means of Cost Over-run in New York City with the United States Projects

The above calculation is based on the equation used for hypothesis I. Two sample T-test was used to assess the hypothesis that the average cost over-run in New York City projects is higher than the United States projects. Further, the reduction was more than 5 (7.24-0.62 = 6.62), or the accurate estimate of the model from hypothesis I, so it can be

concluded that increasing each category by *one* Likert score (when less than 5) will yield a reduction in cost overruns.

4.12 Summery

The association between cost over-run and twenty-five sub- principle were assessed. The data comes from responses to a questionnaire used in a survey investigating the respondents' familiarity with the considered lean sub-principles. Since the Likert scale questions belong to subcategories of lean principles, these variables can be highly correlated. Due to this multicollinearity, the simple linear regression assumptions are supposed to fail (Blackburn,2020), and hence, the partial least squares (PLS) method was used.

The model used for both hypotheses was highly significant with $p<0.001$. But the R^2 values were around 50%, which is not very satisfactory. It could be because even when there is high variability in the response variable (cost over-run), adequate data points allow observing a significant association (Mangiafico,2016). Due to the scope of the research, only twenty-five lean sub-principles considered, and there might be various other unaccounted factors affecting the variability of the cost over-run, which can result in a low R^2 value (Nau 2020).

Also, as seen from the substantial difference in R^2 and predicted R^2, their models seem to overfit the data, accounting for unnecessary randomness into the model. It can also result in a low R^2 value (Frost 2017). However, yet can make meaningful conclusions on the significantly associated variables with cost over-run from these models. However, it

will affect the model's predictive accuracy and result in wide prediction intervals (Nau 2020). Hence, regardless of the low R^2 These models can help explain the data significantly rather than get the response variable's average and identify the significant variables (Mangiafica 2016). Furthermore, the diagnostic plots show that the models fit the data well without any anomalies. For more accuracy in predicting future studies, the models can be optimized by selecting only the most significant variables. However, for the present study, all the variables were used to assess the study's objectives.

4.13 The Research Product (LPCM Model)

The study results showed which Lean Construction Principles (LCPs) critical factors have significantly influenced average cost overruns in construction projects. The most significant factors that include *eight* principles and *ten* sub-principles formed a new model called Lean Principles in Construction Management (LPCM). By applying this model as part of the Project Management Plan (PMP), engineers and project managers will experience less cost overruns. The LPCM model is shown in table 4.13.

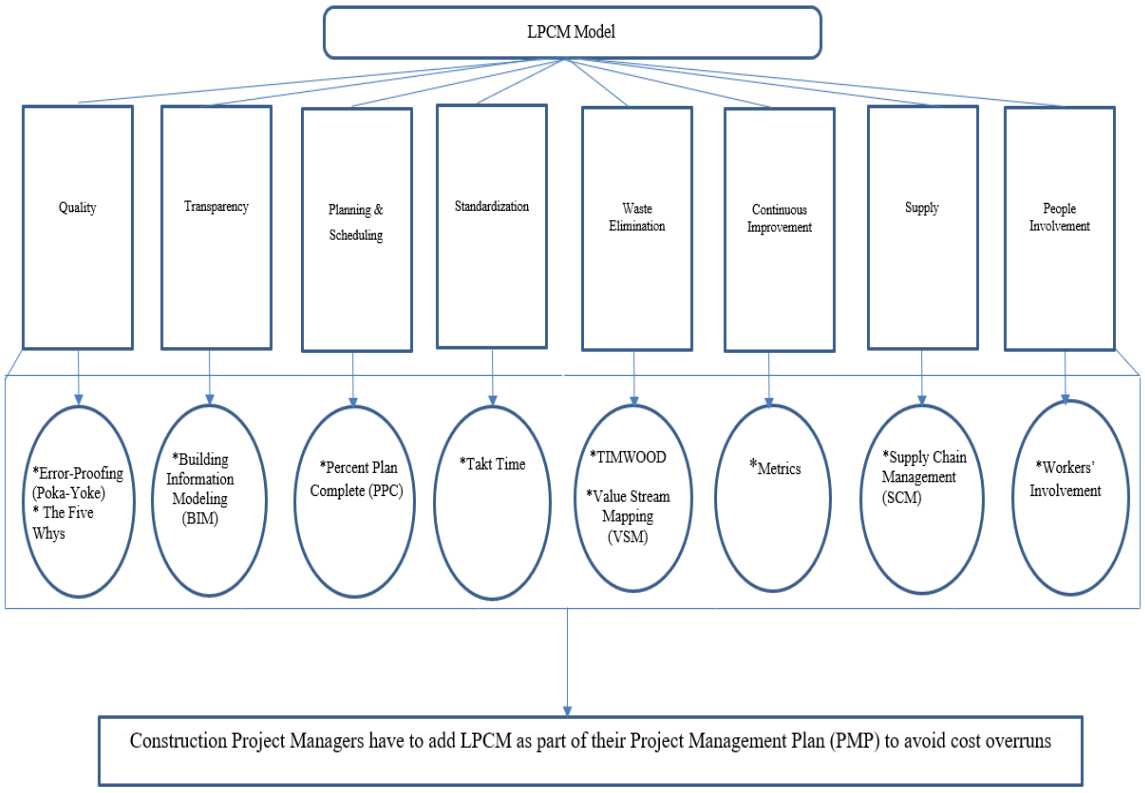


Figure 4.13 Lean Principles in Construction Management (LPCM)

Chapter 5—Discussion and Conclusions

5.1 Discussion

This chapter discusses the findings of the study. Firstly, it speaks to the number of Lean Construction Principles (LCPs) identified and used in the study. Although various studies have identified tens of lean construction principles that can address the problem of inefficiency in the construction industry, this study identified nine major principles that can address cost overruns in the construction industry.

The decrease in the construction industry's number of principles relative to the manufacturing industry did not surprise. It is important to remember that the first time the lean concept was utilized in the manufacturing industry was in the 1950s; the construction industry became familiar with it a few decades later (Jamil and Fathi, 2016). According to studies, the slow pace of implementing lean construction principles in the construction industry is fragmented (Hame, Kowang, and Fei, 2017).

The study results indicated that out of the twenty-five sub- principles of lean construction, *TIMWOOD* emerged as the leading sub-principle of perceived importance, followed by *Supply Chain Management (SCM)* and *Metrics.* This implies that project managers perceive *waste elimination* as the essential construction industry's most crucial lean construction principle. A summary of the ranking of the most significant sub-principles (top ten) is provided in table 5.1 below:

Principles	*Sub-principles*	*Rank*
1. Waste Elimination	TIMWOOD	1
2. Supply	Supply Chain Management (SCM)	2
3. Continuous Improvement	Metrics	3
4. Transparency	Building Information Modeling (BIM)	4
5. People Involvement	Workers' Involvement	5
6. Quality	Error-Proofing (Poka-Yoke)	6
7. Planning and Scheduling	Percent Plan Complete (PPC)	7
8. Quality	The Five Whys	8
9. Waste Elimination	Value Stream Mapping (VSM)	9
10. Standardization	Takt Time	10

Table 5.1 Ranking of the Most Significant Factors

This study has established that many project managers in the construction industry are aware of the existing lean construction subprinciples and have sometimes used them. However, these sub-principles that serve as critical tools for implementing lean construction principles are not used consistently to reduce cost overruns in the construction industry. Similarly, the study's findings indicate that some project managers have never applied lean construction principles in their construction projects, hence cost overruns in their construction projects continue to rise.

5.2 Conclusions

The most significant factors that include *eight* principles and *ten* sub-principles formed a new model called Lean Principles in Construction Management (LPCM). The study established that the LPCM could reduce cost overruns. The results obtained in this study using a machine learning methodology provided strong statistically substantiated support to research hypotheses. In particular, it has been shown that:

I. Currently, the average cost overruns in New York City (7.2%) compared to the nationwide construction project (5.6%) are higher.

II. The results obtained in this study using a machine learning methodology provided statistically substantiated support to research hypothesis I and research hypothesis II.

III. Application of the Lean Principles in construction projects in NYC (measured by 1 unit increase in the relevant Likert scores (up to 5) of the variables in the model) will result in an average change of cost overruns (when other variables are held constant).

The study has established that the problem of cost overruns persists in the construction industry, and there is a need to develop an effective strategy towards its mitigation. To effectively address the cost overruns, the study sought to determine the critical causes of cost overruns in the construction industry. An extensive literature review indicated that various studies had identified numerous causes of cost overruns in construction projects. The study by Bajjou & Chafi (2018) established that apart from the traditional causes of cost overruns in construction projects, waste is one of the significant causes of overruns in the construction industry. The study identified seven critical causes of waste in the construction industry.

Through an extensive literature review, the study identified nine fundamental principles and twenty-five sub-principles of lean construction that can be used to mitigate cost overruns in the construction industry. The Lean Construction Principles include Customer Focus, Continuous Improvement, People Involvement, Quality, Standardization, Supply, Planning and Scheduling, Waste Elimination, and Transparency.

These Lean Construction Principles need to be implemented using the twenty-five sub-principles that can be used to control various activities in the construction process in a manner that helps to reduce cost overruns.

This study recommends applying the Lean Principles in Construction Management (LPCM) as an effective tool that helps construction project managers and engineers reduce cost overruns and improve the overall project performance by eliminating the intangible elements of waste. Unlike other studies that recommend lean construction methods, practices, and techniques to improve project performance in the construction industry, this study has a practical lean construction principles model that engineers and project managers can add to their Project Management Plan (PMP) to avoid cost overruns.

5.3 Limitations of the Study

The study adopted a quantitative approach, which focuses on the cause-effect relationship between the cost overruns and lean construction principles. Therefore, the study fails to consider the underlying issues, beliefs, and cultural practices that lead to cost overruns in the construction industry, as a qualitative research approach could do. The positivism applied in the quantitative survey cannot be relied upon to explain and interpret the actions of project managers that contribute to cost overruns in the construction industry but instead focuses on the known causes of cost overruns, which is an iceberg of the more significant problem of cost overruns in New York City capital projects.

An online survey may lead to wrong and misleading responses. Since there is a lack of accountability, participants' chances to misrepresent or give misleading figures to

finish giving responses are very high. Participants may not accurately contribute to the current study (Vehovar & Katja, 2008).

5.4 Contributions to Body of Knowledge

The current literature body comprises lean tools, techniques, concepts, approaches, and strategies used independently to achieve various construction industry objectives.

The LPCM is an innovative and effective way of addressing cost overruns as it can be applied alongside the existing tools, techniques, and methods to ensure the gains made towards reducing cost overruns in the construction industry. The study builds a deeper understanding of waste as a significant cause of cost overruns in the construction industry. Wastage is currently understood in real terms, primarily material wastage during the construction process (Bajjou and Chafi,2018). This study explored the concept of waste from a different perspective.

The study will benefit both scholars and practitioners as it provides general guidelines on the identified lean principles and tools that lead to a better understanding of the lean construction philosophy.

5.5 Recommendations for Future Research

The research recommends further investigations into the effectiveness of the proposed Lean Principles in Construction Management (LPCM) using a qualitative study that will focus on the shortcomings, challenges, and obstacles that may hinder the application of lean construction principles in construction projects.

This research recommends further investigation into the reasons for the poor implementation of lean construction in the construction industry in order for effective measures to be put in place to improve the rate of use of lean construction principles in construction projects.

The study recommends detailed studies using quantitative and qualitative approaches to identify more lean construction principles and sub-principles in the construction industry.

The study also suggests stratification factors such as industry type, countries, region, cities, budgeting and involve a broader range of construction industries to build the information accessible for future correlation among different waste factors and affirm the overall unwavering quality and legitimacy of the outcomes acquired from the current study.

Last but not least, this research recommends further study into the causes of schedule delays and finding solutions in the construction industry.

References

Ajayi S.O., Oyedele O.L., Akinade O.O., MuhammadB H.A. Alaka, Owolabi H.A., Kadiri K.O. (2017). Attributes of design for construction waste minimization: a case study of waste-to-energy projects, Renew. Sustain. Energy Rev. 73, 1333–1341.

Aziz, R. F., & Hafez, S. M. (2013). Applying lean thinking in construction and performance improvement. Alexandria Engineering Journal, Faculty of Engineering, Alexandria University, 52(4), 679–695.

Babalola O., Ibem E.O. & Ezema C.I. Implementation of lean practices in the construction industry: A systematic review. Building and Environment 148 (2019) 34–43.

Bajjou S. Mohamed, Cha Anas & Ennadi Abdelali, Development of a Conceptual Framework of Lean Construction Principles: An Input-Output Model. Journal of Advanced Manufacturing Systems Vol. 18, No. 1 (2019) 1–34.

Blaikie, N. (2007). Approaches to social inquiry, 2nd ed. Cambridge: Polity Press.

Denzin, N. K., & Lincoln, Y. S. (1998). The landscape of qualitative research: Theories and issues. London: SAGE Publications.

Emuze, F. A., & Saurin, T. A. (Eds.). (2015). Value and waste in lean construction. London, UK: Routledge.

Haddad, M. G., Zouein, P. P., Salem, J., & Otayek, R. (2016). Case study of lean in hospital admissions to inspire culture change. Engineering Management Journal, 28(4), 209–223.

Johansen, E., Walter L. (2007). Lean construction: prospects for the German construction industry, Lean Construction Journal 3 (1) (2007) 19–32.

Khanh, H. D., & Kim, S. Y. (2014a). Identifying causes for waste factors in high-rise building projects: A survey in Vietnam. KSCE Journal of Civil Engineering, 18(4),865–874.

Minami, N. A., Soto, L. L., & Rhodes, D. H. (2010). Dynamic lean management of the naval construction process. Engineering Management Journal, 22(2), 36–43.

Monden Y. (1998). Toyota production system: An integrated approach to just-in-time, 3rd Ed., Engineering and Management, Norcross, Georgia.

Ong, S. E (2003). Mass spectrometric-based approaches in quantitative proteomics. Methods, 29(2), pp.124–130.

Saunders, M., Lewis, P. & Thornhill, A. (2009). Research Methods for Business Students 5th ed., Essex, England: Pearson Education Limited.

Tezel A. & Nielsen Y. (2013). Lean Construction Conformance among Construction Contractors in Turkey. Journal of Management in Engineering. DOI: 10.1061/(ASCE)ME.1943-5479.0000145

Lapinski, A.R., Horman, M.J. & Riley, D.R. (2006). Lean Processes for Sustainable Project Delivery. *Journal of Construction Engineering and Management*, vol. 132 no. 10, pp. 1083-1091.

Abdulmalek, F. & Rajgopal, J. (2007). Analyzing the benefits of lean manufacturing and value stream mapping via simulation: A process sector case study, International Journal of Production Economics 107(1), 223–236.

Adedeji B. Badiru. (2018). Quality Management in Construction Projects, 2nd Ed. Industrial Innovation Series. Air Force Institute of Technology (AFIT)—Dayton, Ohio.

Al Hattab, M. & Hamzeh, F. (2015). Using social network theory and simulation to compare traditional versus BIM–lean practice for design error management, Automation in Construction, Elsevier, 52, 59–69

Al-Aomar, R. (2012). Analysis of lean construction practices at the Abu Dhabi construction industry. Lean Construction Journal, 105–121.

Aneesa M., Gupta A.K & Desa D.B (2015). Supply Chain Management: Effective Tool in the Construction Industry. International Journal of Novel Research in Engineering and Science Vol. 2, Issue 1, pp, 35-40.

Aziz, F. & Hafez. S. (2013) Applying lean thinking in construction and performance improvement, Alexandria Engineering Journal, 52 (4), 679–695.

Bajjou, M. & Chafi, A. (2018). A conceptual model of lean construction: A theoretical framework. Malaysian Construction Research Journal, 26(3), 67-86.

Bajjou, M., Chafi, A. & Ennadi, A. (2019). Development of a Conceptual Framework of Lean Construction Principles: An Input-Output Model. Journal of Advanced Manufacturing Systems, 18(1), 1–34.

Bajjou, M., Chafi, A., Ennadi, A., and Hammoumi, M. (2017). The practical relationships between lean construction tools and sustainable development: A literature review, J. Engineering, Science and Technology Review, 10, 170–177.

Ballard G. Can Pull Techniques Be Used In Design Management? Presented at the Conference on Concurrent Engineering in Construction, Helsinki, Finland. August 26-27, 1999.

Bemelmans J., Voordijk H., Vos B. & Buter J. (2012). Assessing Buyer-Supplier Relationship Management: Multiple Case-Study in the Dutch Construction Industry: Case Study. Journal of Construction Engineering and Management, Vol. 138, No. 1.

Briscoe, G. & Dainty, A. R. J. (2005). Construction supply chain integration: An elusive goal? Supply Chain Management International journal 10(4), 319–326.

Bygballe, E. & Ingemansson, M. (2014). The logic of innovation in construction, Industrial Marketing Management, 43(3), 512–524.

Cui, H. (2014). Value theory of lean construction in China. Advanced Materials Research, 955–959, 4138–4141.

Devaki, M. & Jayanthi, R. (2014). Barriers to implementation of lean principles, International Journal of Engineering Research and Technology, 3, 1189–1192.

Diekmann J, Balonick J, Krewedl M, Troendle L. (2003). Measuring lean conformance. In Proc. 11th Annual Conference of the International Group for Lean Construction, 2–8.

Diekmann, J., Krewedl, M., Balonick, J., Stewart, T. & Won, S. (2004). Application of lean manufacturing principles to construction. Construction Industry Institute, 191, 1–325.

Elmaghraby, S. E., Herroelen, W. S., & Leus, R. (2003). Note on the paper 'Resource-constrained project management using enhanced theory of constraint' by Wei et al. International Journal of Project Management, 21(4), 301–305.

Forcada N., Gangolells M., Casals M. & Macarulla M. (2017). Factors affecting rework costs in construction, Journal Construction Engineering Management, 143 (8).

Formoso, C., Dos. Santos, A., and Powell, J. (2002). An exploratory study on the applicability of process transparency in construction sites, Journal of Construction Research, 3, 35–54.

Frandson, A., Berghede, K., & Tommelein, I. D. (2014, June). Takt-time planning and the last planner. In Proc. 22nd Ann. Conf. of the Int'l Group for Lean Construction. Group for Lean Construction, 23-27.

Gudehus, T. & Kotzab, H. (2009). Comprehensive logistics. Heidelberg: Springer Berlin Heidelberg. Doi: 10.1007/978-3-540-68652-1

Harker, A., Allcorn, W., & Taylor, D. (2007). Material logistic plan. Good practice guidance. Bunbury, OX.

Harris, F. & McCaffer, R. (2013). Modern Construction Management. Construction Management and Economics, 31(12), 1215-1217. doi:10.1515/9783990434550

Hook, M., Stehn, L., 2008. Lean Principles in Industrialized Housing Production: The Need for a Cultural Change Lean Culture. Lean Construction Journal, 20–33

Huhnt, W. (2005). Generating sequences of construction tasks. In Proceedings of 22nd of W78 conference on information technology in construction, Dresden, Germany, 17-22.

Ikuma, L. Nahmens, I., and James, J. (2011). Use of safety and lean integrated Kaizen to improve performance in modular home building, Journal of Construction Engineering Management, 137, 551–560.

Issa, U.H. (2013) Implementation of lean construction techniques for minimizing the risks effect on project construction time. Alexandria Engineering Journal, 52, 697–704. doi:10.1016/j.aej.2013.07.003

Jørgensen, B & Emmitt, S. (2008). Lost in transition: the transfer of lean manufacturing to construction. Engineering, Construction and Architectural Management, 15(4), pp.383-398

Kerzner, H. (2009). Project management: A systems approach to planning, scheduling, and controlling, 11th ed. Hoboken, NJ: John Wiley & Sons.

Khanh, H. & Kim, S. (2014). Identifying causes for waste factors in high-rise building projects: A survey in Vietnam, KSCE Journal of Civil Engineering, 18, 865–874.

Kim, D. and Park, H. Innovative construction management method: Assessment of lean construction implementation, KSCE Journal of Civil Engineering, 10, 381–388.

Lai, K.h. & Cheng, T. C. E. (2009). Just-in-time logistics. Burlington, VT: Gower.

Low Sui, P. & Gao, S. (2011). The Application of the Just-in-Time Philosophy in the Chinese Construction Industry. Journal of Construction in Developing Countries, 16, 91-111.

Mallawaarachchi H. & Senaratne S.Importance of quality for construction project success, International Conference on Structural Engineering and Construction Management, Kandy, Sri Lanka, 2015.

Mitchell, A., Frame, I., Coday, A. & Hoxley, M. (2011). A conceptual framework of the interface between the design and construction processes. Engineering, Construction, and Architectural Management, 18(3), pp. 297-311.

Mossman, A. (2009). Creating value: A sufficient way to eliminate waste in lean design and lean production, Lean Construction Journal, 13–23.

Natasya, A., Wahab, A., Mukhtar, M., Sulaiman, R. (2013). A Conceptual Model of Lean Manufacturing Dimensions. Procedia Technology, 11, 1292–1298. doi:10.1016/j.protcy.2013.12.327

Nesensohn C., S. T. Demir, and D. J. Bryde Developing the True North route map as a navigational compass in a construction project management organisation, Lean Constr. J. (2013) 1–18.

Nesensohn, C., Bryde, D., and C. Pasquire (2015). A measurement model for lean construction maturity, inProc. IGLC 2323rd Annual Conference International Group for Lean Construction Global Knowledge Solution, vol. 6, pp. 652–660

Nowotarski, P. & Milwicz, R. (2015). Management of electrical modernization in construction with the comparison between standard and flexible approach in the economic aspects, Procedia Engineering, 122, 196–203.

Rajan K., Dwivedi R.K. & Ajay V. (2016). Poka-yoke technique, methodology & design. Indian Journal of Engineering, 13(33), 362-370.

Sacks R., Treckmann M. & Rozenfeld O. (2009). Visualization of work how to support lean construction, Journal of Construction Engineering Management 135, 1307–1315.

Sacks, R., Treckmann, M., Rozenfeld, O. (2009). Visualization of Work Flow to Support Lean Construction. Journal of Construction Engineering and Management, 135, 1307–1315. doi:10.1061/(ASCE)CO.1943-7862.0000102

Salem, O., Solomon, J., Genaidy, A. & Luegring, M. (2005). Site implementation and assessment of lean construction techniques, Lean Construction Journal, Vol.2, pp.2–21.

Salem, O., Solomon, J., Genaidy, A. & Minkarah, I. (2006). Lean construction: From theory to implementation, Journal of Management Engineering, 22, 168–175.

Savolainen, J. Kahkonen, K. Niemi, O. Poutanen, J., and E. Varis, E. (2015). Stirring the construction project management with co-creation and continuous improvement, Procedia Economics Finance, 21, 64–71.

Tauriainen, M. Marttinen, P. Dave, B., and Koskela, L. (2016). The effects of BIM and lean construction on design management practices, Procedia Engineering, 164, 567–574.

Tezel, A. & Nielsen, Y. (2013). Lean construction conformance among construction contractors in Turkey, Journal of Management Engineering, 29(3), 236–250.

Tezel, B.A., Koskela, L.J., Tzortzopoulos, P., Formoso, C.T., Alves, T., Neto, B., Viana, D., Mota, B. (2010). Process transparency on construction sites: Examples from construction companies in Brazil. Proc. 18th Annual Conference International Group for Lean Construction, 296–305.

Tiwaril, S., Magar, R. & Honnutagi, A. (2018). Resource Optimization for Sustainable Construction: A State of Art. International Advanced Research Journal in Science, Engineering and Technology, 5(3), 69-73.

Tommelein D.I. & Demirkesen S. (2018). Mistake proofing the Design of Construction Processes Using Inventive Problem Solving (TRIZ). CPWR-Report.

Trivedi, J. & R. Kumar, R. (2014). Optimisation of construction resources using lean construction technique, International Journal of Engineering Management and Economics, 4(3–4), 213–228.

Ullah F., Thaheem M. J., Siddiqui S. Q. and Khurshid M. B. (2017). Influence of six sigma on project success in the construction industry of Pakistan, Total Quality Management Journal, 29, 276–309.

Vrijhoef, R. & Koskela, L. (2000). The four roles of supply chain management in construction." European Journal Purchase and Supply Management 6(3–4), 169–178.

Wang, J., Ma, R.Q. (2014). Based on Lean Construction Standardization Management Analysis of the Performance Evaluation Index System. Advanced Materials Research, 933, 879–883. doi: 10.4028/www.scientific.net/AMR.933.879

Yassine, T., Bacha, M. B. S., Fayek, F., & Hamzeh, F. (2014). Implementing takt-time planning in construction to improve workflow. In *Proc. 22nd Annual Conference of the International Group for Lean Construction*, 787-798.

Yin, Y., Qin, S. & Holland, R. (2011). Development of a design performance measurement matrix for improving collaborative design during a design process. International Journal of Productivity and Performance Management, 60(2), pp. 152-184.

Zeynalian, M., Trigunarsyah, B., & Ronagh, H. (2013). Modi¯cation of advanced programmatic risk analysis and management model for the whole project life cycle's risk. Journal of Construction Engineering Management, 138, 51–60.

Aliaga, M. and Gunderson, B. (2002) Interactive Statistics. [Thousand Oaks]: Sage Publications.

Creswell, J. W. (2009). Research Design Qualitative, Quantitative, and Mixed Methods Approach, 3rd ed. London: SAGE Publication.

Creswell, J. W. (2013). Qualitative inquiry & research design 3rd Ed., Thousand Oaks, CA: Sage Publications.

Evans R.J. & Mathur A. (2005). The value of online surveys. Internet Research Vol. 15 No. 2, pp. 195-219.

Fleetwood, S. (2005). Ontology in organization and management studies: A critical realist perspective, Organization, Vol. 12, pp. 197–222.

Gill, J., and Johnson, P. (2010). Research Methods for Managers, 4th Ed. London: Sage.

Holden, M. T., & Lynch, P. (2004). Choosing the appropriate methodology: Understanding research philosophy. The Marketing Review, 4(4), 397-409.

Johnson, B. & Christensen L. (2008). Educational research: Quantitative, qualitative, and mixed approaches. Thousand Oaks, CA: Sage Publications.

Kim T. Kyun (2015). T-test as a parametric statistic. Korean Journal of Anesthesiology, pp 541-546.

Majid U. (2018). Research Fundamentals: Study Design, Population, and Sample Size, URNCST Journal, Volume 2, Issue 1.

Nigel Mathers, Nick Fox & Amanda Hunn (2009). Surveys and Questionnaires, the NIHR Research Design Service for Yorkshire & the Humber.

Saunders M., Lewis P. & Thornhill A. (2015). Understanding research philosophy and approaches to theory development. Pearson Education Limited, England.

Saunders, M., Lewis, P. & Thornhill, A. (2012) "Research Methods for Business Students" 6th edition, Pearson Education Limited p.288

Sharma, G. (2017). Pros and cons of different sampling techniques. International journal of applied research, 3(7), 749-752.

Simpson, B. (2009). Pragmatism, Mead, and practice turn. Organization Studies, 30(12), 1329-1347.

Sullivan M. Gail & Artino R. Anthony (2013). Analysing and Interpreting Data from Likert-type Scales, Journal of Graduate Medical Education, 5(4): 541–542.

Taherdoost, H. (2016). Sampling Methods in Research Methodology; How to Choose a Sampling Technique for Research. International Journal of Academic Research in Management (IJARM) Vol. 5, No. 2, Page: 18-27.

Uyanik K. G. Guler N. (2013). A study on multiple linear regression analysis. 4th International Conference on New Horizons in Education. Procedia-Social and Behavioral Sciences 106, 234-240.

Van den Broeck J, Sandøy IF, Brestoff JR. (2013). The Recruitment, Sampling, and Enrollment Plan. In Epidemiology: Principles and Practical Guidelines, pp. 171-196. Springer Netherlands.

Vehovar V, & Katja, M. (2008) Overview: Online Surveys. In: The SAGE Handbook of Online Research Methods [Internet], p. 177–94.

Williams, C. (2011). Research methods. Journal of Business & Economics Research (JBER), 5(3).

Wilson, A. & Laskey, N. (2003). Internet-based marketing research: a serious alternative to traditional research methods? Marketing Intelligence & Planning, Vol. 21 No. 2, pp. 79-84.

Tony (June 4, 2020). TIMWOOD: ToughNickel. The Seven Wastes of Lean Manufacturing. Retrieved from https://toughnickel.com/business/Seven-Wastes

El-Namrouty A. K. & AbuShaaban S. M. (2013). Seven wastes elimination targeted by lean manufacturing case study "gaza strip manufacturing firms. International Journal of Economics, Finance and Management Sciences, 1(2): 68-80

Al Tabtabai H.M. (2002). Causes for delays in construction projects in Kuwait. Engineering Journal of the University of Qatar,15,19-37.

Kometa S. T., Olomolaive P. O., & Harris, F. C. (1994). Attributes of UK construction clients influencing project consultants' performance. *Construction Management and economics,*12(15),433-443.

Bajjou, M. & Chafi, A. (2018). The potential effectiveness of lean construction principles in reducing construction process waste: An input-output model. Journal of Mechanical Engineering and Sciences, 12(4), 4141-4160.

Frandson Adam (2019). Takt Time planning as a Work Structuring Method to Improve Construction Workflow. A dissertation submitted in partial satisfaction of the requirements for the Degree of Doctor of Philosophy in Engineering –Civil and Environmental Engineering, University of California, Berkeley.

Hame N.B.M. Kowang T.O. & Fei G.C. (2017). Categorization of lean research and development tools and techniques: a process-based approach, *Indian Journal of Science and Technology* 10 (3), 1–7.

Sarhan, J. G., Xia, B., Fawzia, S., & Karim, A. (2017). Lean construction implementation in the Saudi Arabian construction industry. *Construction Economics and Building*, 17(1), 46–69

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis: A global perspective* (8th edition). United Kingdom: Cengage Learning EMEA.

Malhotra, N. K. (2010). *Marketing research: An applied orientation* (6th edition). New Jersey: Prentice-Hall.

George, D (2011), SPSS for Windows Step by Step: A Simple Study Guide and Reference, 17.0 Update, 10/E, Pearson Education India.

Nunnally, J. C. (1978). Psychometric Theory. (2nd ed.). New York: McGraw-Hill.

Bartels, R. (1982). The Rank Version of von Neumann's Ratio Test for Randomness. *Journal of the American Statistical Association* 77(377), 40-46

Oluwatosin Babalola, Eziyi O. Ibem, Isidore C.Ezema. "Implementation of lean practices in the construction industry: A systematic review", Building and Environment, 2019

Mohamed S. Bajjou, Anas Chafi, Abdelali Ennadi. "Development of a Conceptual Framework of Lean Construction Principles: An Input-Output Model" Journal of Advanced Manufacturing Systems, 2019

Ning, S.P. "Genome composition and genetic diversity of Musa germplasm from China revealed by PCR-RFLP and SSR markers", Scientia Horticulturae, 20071120

Michael J. Crawley. "The R Book", Wiley, 2012

Wallace Avelar, Marcelo Meiriño, Guilherme Luiz Tortorella. "The practical relationship between continuous flow and lean construction in SMEs", The TQM Journal, 2019

Bassam A. Tayeh, Wesam Salah Alaloul, Noor K. Al-Ghazalli. "Tender Pricing of Infrastructure Projects: Affecting Factors", International Conference on Sustainable Infrastructure 2019

Li-Ting Chen, Ya-Wen Hsu. "Socio-Ecological Predictors of Frequent Bike Share Trips: Do Purposes Matter?", International Journal of Environmental Research and Public Health, 2020

Sandholm, William. "Vital Statistics", Oxford University Press

Bhim Singh, Suresh K. Garg, Surrender K.Sharma. "Value stream mapping: literature review and implications for Indian industry", The International Journal of Advanced Manufacturing Technology, 2010

Zakaria Dakhli, Zoubeir Lafhaj, Antoine Bos." Experiencing Lean Six Sigma in the French residential construction", International Journal of Lean Six Sigma, 2016

A M Faten Albtoush, S I Doh, R A Rahman." Underlying factors of cost overruns in developing countries: multivariate analysis of Jordanian projects", IOP Conference Series: Earth and Environmental Science, 2021

Anil Chaudhary, Ashish Kumar Singh, M.L.Meena. "Productivity improvement of an electrical appliance industry by implementing lean manufacturing tools and a low-

cost intervention (a case study)", International Journal of Productivity and Quality Management, 2020

Jussi Savolainen, Kalle Kähkönen, Olli Niemi, Jenni Poutanen, Elisa Varis. "Stirring the Construction Project Management with Cocreation and Continuous Improvement", Procedia Economics and Finance, 2015

Xi Jiang, Jorge Gonzalez-Martinez, Eric Halgren. "Posterior hippocampal spindle-ripples co-occur with neocortical theta-bursts and down-upstates, and phase-lock with parietal spindles during NREM sleep in humans", Cold Spring Harbor Laboratory, 2019

"Digital Conversion on the Way to Industry 4.0", Springer Science and Business Media LLC,2021

"Proceedings of the 21st International Symposium on Advancement of Construction Management and Real Estate", Springer Science and Business Media LLC, 2018

Conesa, A. "Direct functional assessment of the composite phenotype through multivariate projection strategies", Genomics, 200812

Ghaleb J. Sweis, Mohammad Hiyassat, Fares F.Al-Hroub. "Assessing lean conformance by firstgrade contractors in the Jordanian construction industry", Construction Innovation, 2016

Martin Evans, Peter Farrell. "Barriers to integrating building information modelling (BIM) and lean construction practices on construction mega-projects: a Delphi study", Benchmarking: An International Journal, 2020

"Collaboration and Integration in Construction, Engineering, Management and Technology", Springer Science and Business Media LLC, 2021

Adib Amany, Katayoon Taghizade, Esmatullah Noorzai. "Investigating conflicts of expert contractors using the last planner system in building information modeling process", Journal of Engineering, Design, and Technology, 2020

Avinash Panwar, Bimal Nepal, Rakesh Jain, Ajay P.S. Rathore, Andrew Lyons."Understanding the linkages between lean practices and performance improvements in Indian process industries", Industrial Management & Data Systems, 2017

Jelena M. Andrić, Abdul-Majeed Mahamadu, Jiayuan Wang, Patrick X. W. Zou, Ruoyu Zhong." The Cost Performance and Causes of Overruns in Infrastructure Development Projects in Asia", Journal of Civil Engineering and Management, 2019

Jerusalem Yizengaw. "The Implementation of School Improvement Program in Ethiopia: Emphasis on Bahir Dar City Secondary Schools", SAGE Publications, 2020

Joy Ong, Low Sui Pheng. "Waste Reduction in Precast Construction", Springer Science and Business Media LLC, 2021

Low Sui Pheng, Wu Min. "Just-in-time management in the ready mixed concrete industries of Chongqing, China and Singapore", Construction Management and Economics, 2005

Schimanski, Marcher, Monizza, Matt. "The Last Planner® System and Building Information Modeling in Construction Execution: From an Integrative Review to a Conceptual Model for Integration", Applied Sciences, 2020

Senem Seyis, Esin Ergen, Emilio Pizzi. "Identification of Waste Types and Their RootCauses in Green-Building Project Delivery Process", Journal of Construction Engineering and Management, 2016

Yau-Ren SHIAU, Fang-Yi LO, Po-Cheng KO. "Early intervention mechanism for preventing electrocution in construction engineering", 154 Industrial Health, 2021

Alexandre M. J. -C. Wadoux, Brendan Malone,Budiman Minasny, Mario Fajardo, Alex B.McBratney. "Chapter 8 Selection of the Samples for Laboratory Analysis", Springer Science and Business Media LLC, 2021

"The Construction Industry in the Fourth Industrial Revolution", Springer Science and Business Media LLC, 2020

Hammad AlNasseri, Radhlinah Aulin. "Assessing Understanding of Planning and Scheduling Theory and Practice on Construction Projects", Engineering Management Journal,2015

Shakil Ahmed, Md. Mehrab Hossain, Iffat Haq."Implementation of lean construction in the construction industry in Bangladesh: awareness, benefits and challenges", International Journal of Building Pathology and Adaptation, 2020

Shang Gao, Sui Pheng Low. "Lean Construction Management", Springer Science and Business Media LLC, 2014

The Five Whys Technique O. Serrat, Knowledge Solutions, DOI 10.1007/978-981-10-0983-9_32

Cresswell JW, Plano Clark VL. *Designing and conducting mixed-method research.* 2nd Sage; Thousand Oaks, CA: 2011.

Breiman L. (2001) *Random forests. Machine Learning,*45, 5-32

Mevik, B., R. Wehrens (2020) Introduction to the PLS package

Timothy D. Blackburn (2020). Data Analysis for Engineers and Scientists, Partial Least Squares (PLS).

Bland, J.M., Altman, D.G. Applying the right statistics: analyses of measurement studies. Ultrasound Obstet Gynecol. 2003 Jul;22(1):85-93.

Frost, J. 2017. Five Reasons Why Your R-squared can be Too High, https://statisticsbyjim.com/regression/r-squared-too-high/ Accessed on 03/03/2021.

Mangiafico, S.S. 2016. Summary and Analysis of Extension Program Evaluation in R, version 1.18.8

Nau, R. 2020. Statistical forecasting: notes on regression and time series analysis https://people.duke.edu/~rnau/rsquared.htm, Accessed on 03/03/2021.

(2011) Mean Squared Error. In: Sammut C., Webb G.I. (eds) Encyclopedia of Machine Learning. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-30164-8_528

Appendix A

Appendix A1: Data set

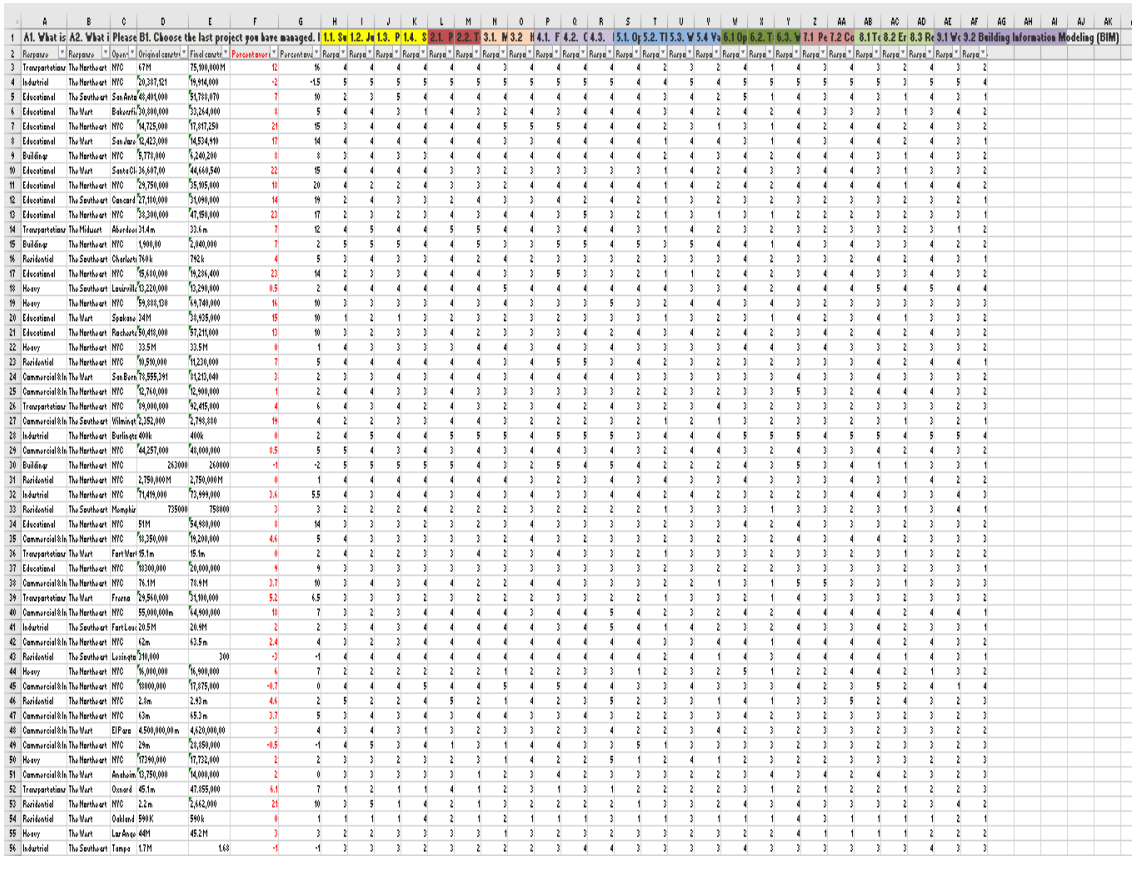


Table1: Data set

Appendix A2: Frequency of Construction Project Managers' Last Project

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Aberdeen	1	.7	.7	.7
	Albany	1	.7	.7	1.3
	Albuquerque	1	.7	.7	2.0
	Anaheim	1	.7	.7	2.6
	Aspen	1	.7	.7	3.3
	Atlanta	1	.7	.7	3.9
	Bakersfield	1	.7	.7	4.6
	Bismarck	1	.7	.7	5.3

	Boise	1	.7	.7	5.9
	Boston	1	.7	.7	6.6
	Burlington	1	.7	.7	7.2
	Charleston	1	.7	.7	7.9
	Charlotte	3	2.0	2.0	9.9
	Chicago	2	1.3	1.3	11.2
	Cincinnati	2	1.3	1.3	12.5
	Cleveland	1	.7	.7	13.2
	Columbus	1	.7	.7	13.8
	Concord	2	1.3	1.3	15.1
	Dallas	1	.7	.7	15.8
	Denver	4	2.6	2.6	18.4
	Detroit	1	.7	.7	19.1
	Durham	1	.7	.7	19.7
	El Paso	1	.7	.7	20.4
	Fort Lauderdale	1	.7	.7	21.1
	Fort Worth	1	.7	.7	21.7
	Fresno	1	.7	.7	22.4
	Greenville	1	.7	.7	23.0
	Houston	3	2.0	2.0	25.0
	Jacksonville	1	.7	.7	25.7
	Kansas City	2	1.3	1.3	27.0
	Lexington	1	.7	.7	27.6
	Lincoln	1	.7	.7	28.3
	Los Angeles	5	3.3	3.3	31.6
	Louisville	1	.7	.7	32.2
	Memphis	1	.7	.7	32.9
	Miami	1	.7	.7	33.6
	Minneapolis	1	.7	.7	34.2
	Nashville	1	.7	.7	34.9
	New Haven	1	.7	.7	35.5
	NYC	63	41.4	41.4	77.0
	Oakland	1	.7	.7	77.6

	OKC	2	1.3	1.3	78.9
	Olympia	1	.7	.7	79.6
	Oregon	1	.7	.7	80.3
	Orlando	1	.7	.7	80.9
	Oxnard	1	.7	.7	81.6
	Philadelphia	2	1.3	1.3	82.9
	Portland	2	1.3	1.3	84.2
	Princeton	1	.7	.7	84.9
	Reno	1	.7	.7	85.5
	Rochester	1	.7	.7	86.2
	Sacramento	2	1.3	1.3	87.5
	Salt Lake	1	.7	.7	88.2
	San Antonio	2	1.3	1.3	89.5
	San Bernardino	1	.7	.7	90.1
	San Diego	1	.7	.7	90.8
	San Francisco	1	.7	.7	91.4
	San Jose	2	1.3	1.3	92.8
	Santa Clara	1	.7	.7	93.4
	Seattle	2	1.3	1.3	94.7
	San Francisco	1	.7	.7	95.4
	Spokane	2	1.3	1.3	96.7
	Springfield	1	.7	.7	97.4
	Takoma	1	.7	.7	98.0
	Tallahassee	1	.7	.7	98.7
	Tampa	1	.7	.7	99.3
	Wilmington	1	.7	.7	100.0
	Total	152	100.0	100.0	

Table2: Frequency of Construction Project Managers' Last Project

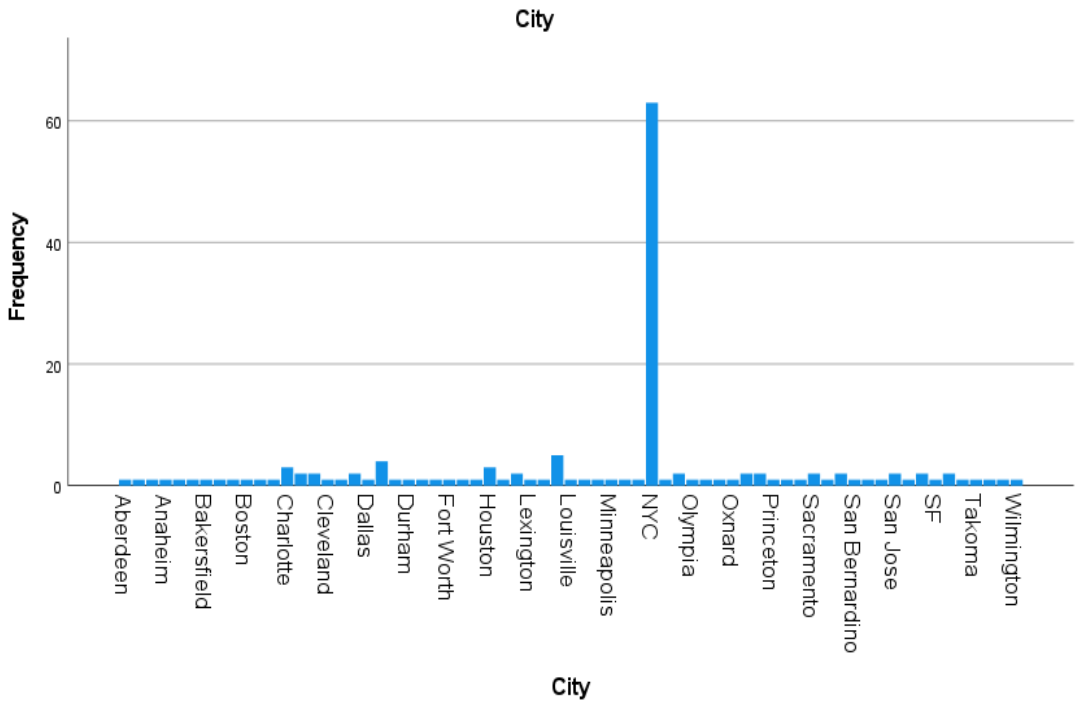
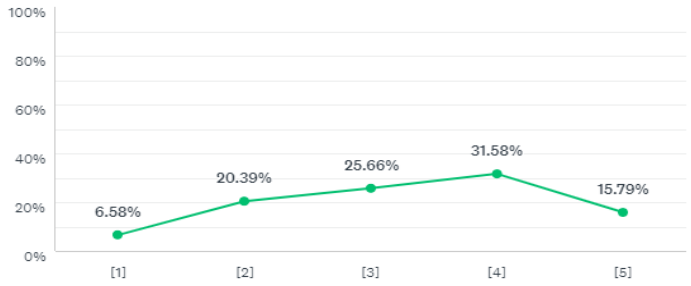


Figure 1: Number of the Data Collected from Construction Project Managers and the Location of their Last project Managed

Appendix A3: Descriptive Statistics

The following figures and tables illustrated the frequency of each of the categories of the Likert scale. The minimum, maximum, median, mean, and standard deviation were shown separately for each sub-principle.

1.1. Supply Chain Management (SCM)

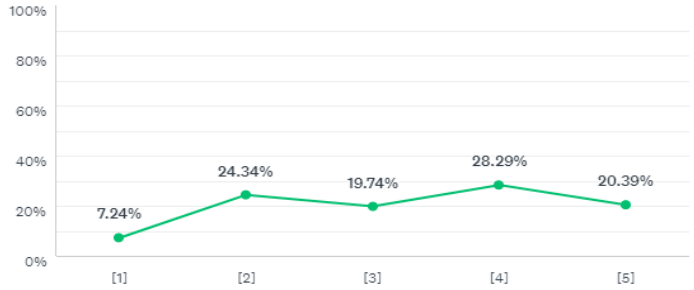


ANSWER CHOICES	RESPONSES	
[1]	6.58%	10
[2]	20.39%	31
[3]	25.66%	39
[4]	31.58%	48
[5]	15.79%	24
TOTAL		152

BASIC STATISTICS

Minimum	Maximum	Median	Mean	Standard Deviation
1.00	5.00	3.00	3.30	1.15

1.2. Just- In - Time (JIT)

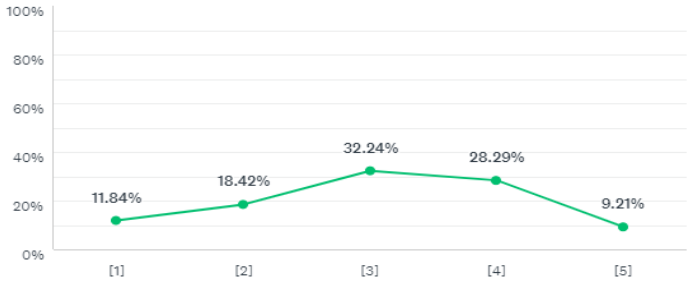


ANSWER CHOICES	RESPONSES	
[1]	7.24%	11
[2]	24.34%	37
[3]	19.74%	30
[4]	28.29%	43
[5]	20.39%	31
TOTAL		152

BASIC STATISTICS

Minimum	Maximum	Median	Mean	Standard Deviation
1.00	5.00	3.00	3.30	1.24

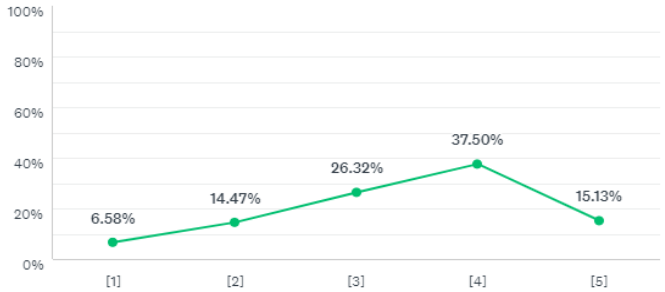
1.3. Pull System



ANSWER CHOICES	RESPONSES	
[1]	11.84%	18
[2]	18.42%	28
[3]	32.24%	49
[4]	28.29%	43
[5]	9.21%	14
TOTAL		152

BASIC STATISTICS				
Minimum	Maximum	Median	Mean	Standard Deviation
1.00	5.00	3.00	3.05	1.14

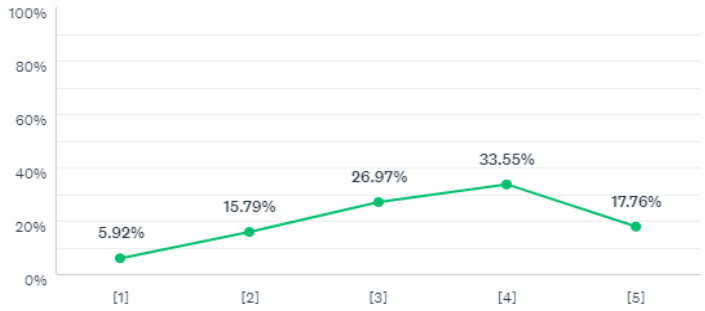
1.4. Suppliers' Involvement



ANSWER CHOICES	RESPONSES	
[1]	6.58%	10
[2]	14.47%	22
[3]	26.32%	40
[4]	37.50%	57
[5]	15.13%	23
TOTAL		152

BASIC STATISTICS				
Minimum	Maximum	Median	Mean	Standard Deviation
1.00	5.00	4.00	3.40	1.11

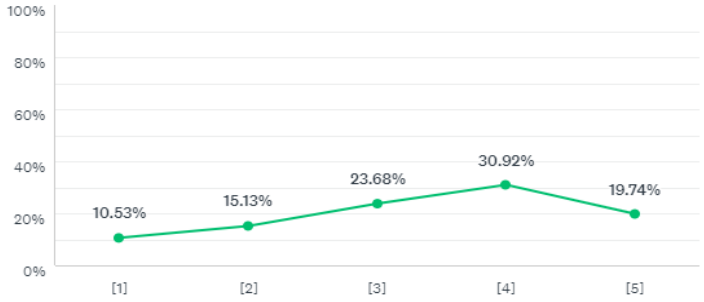
2.1. Workers' Involvement



ANSWER CHOICES	RESPONSES	
[1]	5.92%	9
[2]	15.79%	24
[3]	26.97%	41
[4]	33.55%	51
[5]	17.76%	27
TOTAL		152

BASIC STATISTICS				
Minimum	Maximum	Median	Mean	Standard Deviation
1.00	5.00	4.00	3.41	1.13

2.2. Teamwork



ANSWER CHOICES	RESPONSES	
[1]	10.53%	16
[2]	15.13%	23
[3]	23.68%	36
[4]	30.92%	47
[5]	19.74%	30
TOTAL		152

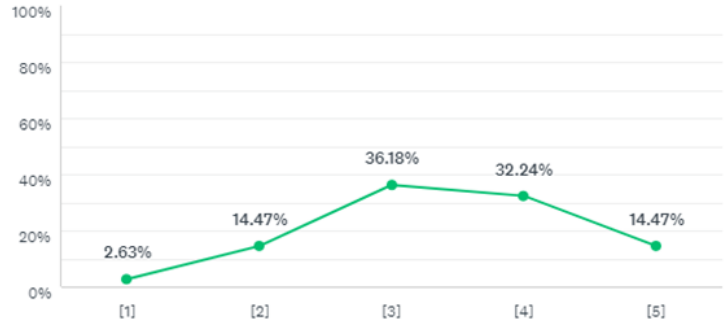
BASIC STATISTICS				
Minimum	Maximum	Median	Mean	Standard Deviation
1.00	5.00	4.00	3.34	1.25

3.1. Metrics

ANSWER CHOICES	RESPONSES	
[1]	8.55%	13
[2]	23.68%	36
[3]	28.29%	43
[4]	23.03%	35
[5]	16.45%	25
TOTAL		152

BASIC STATISTICS				
Minimum	Maximum	Median	Mean	Standard Deviation
1.00	5.00	3.00	3.15	1.20

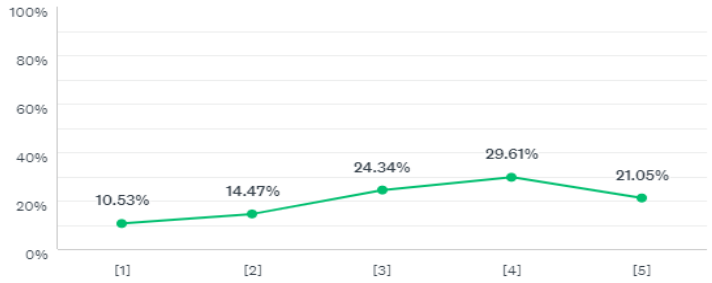
3.2. Huddle Meetings



ANSWER CHOICES	RESPONSES	
[1]	2.63%	4
[2]	14.47%	22
[3]	36.18%	55
[4]	32.24%	49
[5]	14.47%	22
TOTAL		152

BASIC STATISTICS				
Minimum	Maximum	Median	Mean	Standard Deviation
1.00	5.00	3.00	3.41	0.99

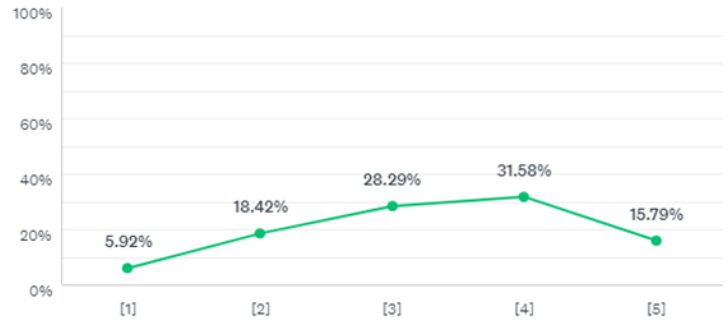
4.1. Flexible Resources



ANSWER CHOICES	RESPONSES	
[1]	10.53%	16
[2]	14.47%	22
[3]	24.34%	37
[4]	29.61%	45
[5]	21.05%	32
TOTAL		152

BASIC STATISTICS				
Minimum	Maximum	Median	Mean	Standard Deviation
1.00	5.00	4.00	3.36	1.25

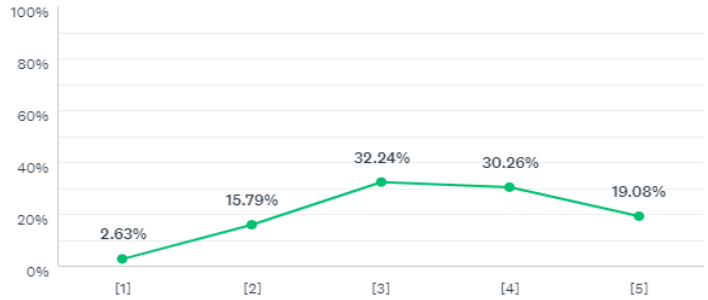
4.2. Optimize Value



ANSWER CHOICES	RESPONSES	
[1]	5.92%	9
[2]	18.42%	28
[3]	28.29%	43
[4]	31.58%	48
[5]	15.79%	24
TOTAL		152

BASIC STATISTICS				
Minimum	Maximum	Median	Mean	Standard Deviation
1.00	5.00	3.00	3.33	1.12

4.3. Relationships

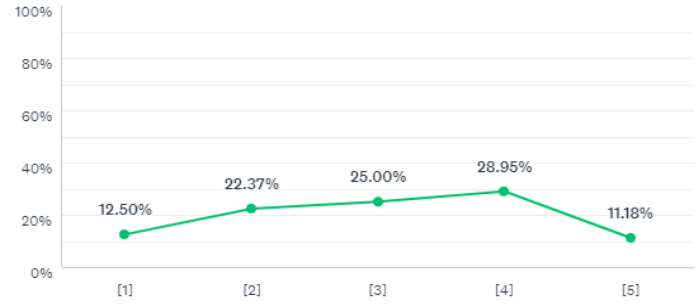


ANSWER CHOICES	RESPONSES	
[1]	2.63%	4
[2]	15.79%	24
[3]	32.24%	49
[4]	30.26%	46
[5]	19.08%	29
TOTAL		152

BASIC STATISTICS

Minimum	Maximum	Median	Mean	Standard Deviation
1.00	5.00	3.00	3.47	1.05

5.1. Optimize Product System

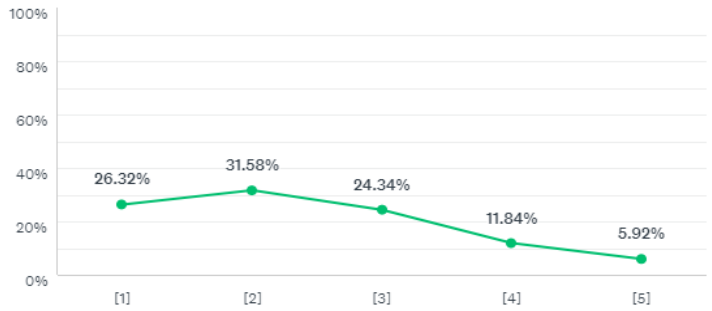


ANSWER CHOICES	RESPONSES	
[1]	12.50%	19
[2]	22.37%	34
[3]	25.00%	38
[4]	28.95%	44
[5]	11.18%	17
TOTAL		152

BASIC STATISTICS

Minimum	Maximum	Median	Mean	Standard Deviation
1.00	5.00	3.00	3.04	1.21

5.2. TIMWOOD

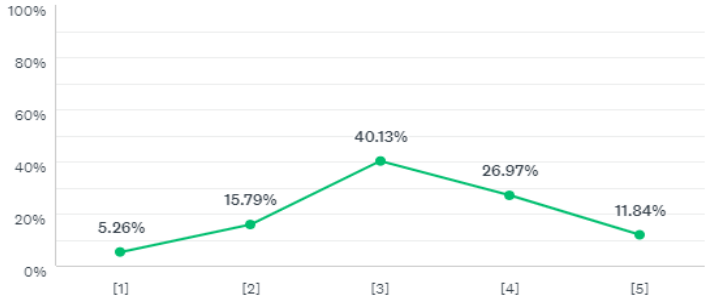


ANSWER CHOICES	RESPONSES	
[1]	26.32%	40
[2]	31.58%	48
[3]	24.34%	37
[4]	11.84%	18
[5]	5.92%	9
TOTAL		152

BASIC STATISTICS

Minimum	Maximum	Median	Mean	Standard Deviation
1.00	5.00	2.00	2.39	1.17

5.3. Waste Disposal Management

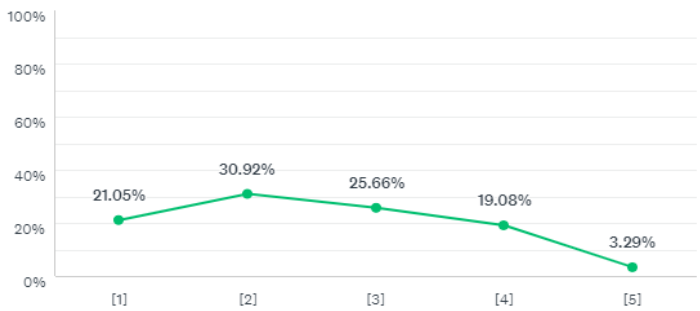


ANSWER CHOICES	RESPONSES	
[1]	5.26%	8
[2]	15.79%	24
[3]	40.13%	61
[4]	26.97%	41
[5]	11.84%	18
TOTAL		152

BASIC STATISTICS

Minimum	Maximum	Median	Mean	Standard Deviation
1.00	5.00	3.00	3.24	1.03

5.4. Value Stream Mapping (VSM)

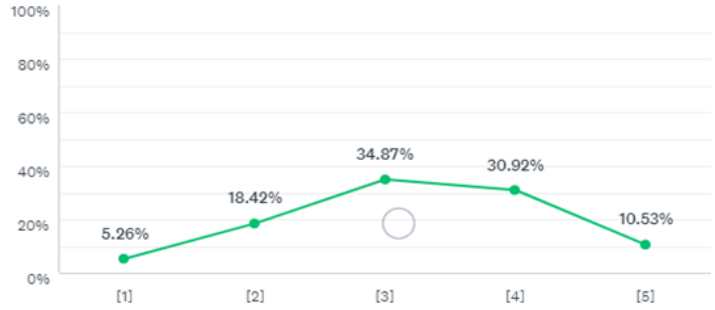


ANSWER CHOICES	RESPONSES	
[1]	21.05%	32
[2]	30.92%	47
[3]	25.66%	39
[4]	19.08%	29
[5]	3.29%	5
TOTAL		152

BASIC STATISTICS

Minimum	Maximum	Median	Mean	Standard Deviation
1.00	5.00	2.00	2.53	1.12

6.1. Optimize Work Content

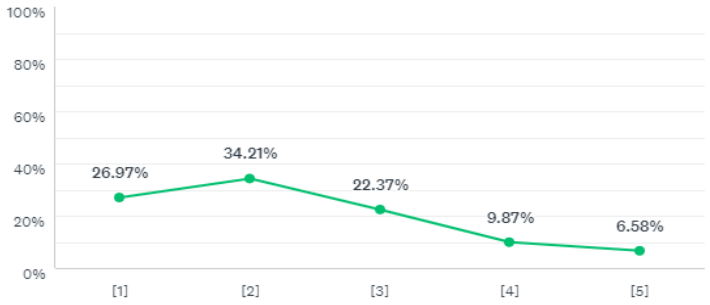


ANSWER CHOICES	RESPONSES	
[1]	5.26%	8
[2]	18.42%	28
[3]	34.87%	53
[4]	30.92%	47
[5]	10.53%	16
TOTAL		152

BASIC STATISTICS

Minimum	Maximum	Median	Mean	Standard Deviation
1.00	5.00	3.00	3.23	1.04

6.2. Takt Time

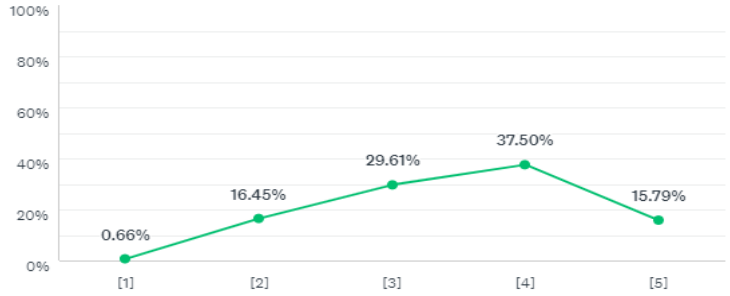


ANSWER CHOICES	RESPONSES	
[1]	26.97%	41
[2]	34.21%	52
[3]	22.37%	34
[4]	9.87%	15
[5]	6.58%	10
TOTAL		152

BASIC STATISTICS

Minimum	Maximum	Median	Mean	Standard Deviation
1.00	5.00	2.00	2.35	1.17

6.3. Work Sequence

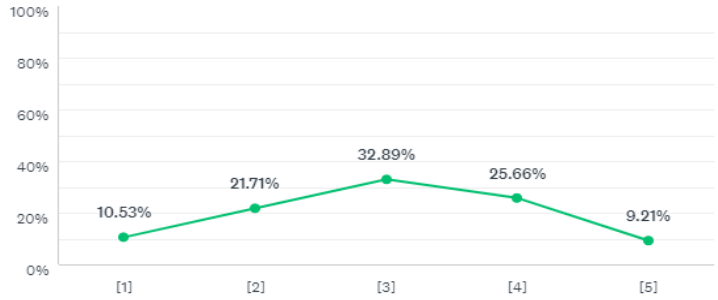


ANSWER CHOICES	RESPONSES	
[1]	0.66%	1
[2]	16.45%	25
[3]	29.61%	45
[4]	37.50%	57
[5]	15.79%	24
TOTAL		152

BASIC STATISTICS

Minimum	Maximum	Median	Mean	Standard Deviation
1.00	5.00	4.00	3.51	0.97

7.1. Percent Plan Complete (PPC)



ANSWER CHOICES	RESPONSES	
[1]	10.53%	16
[2]	21.71%	33
[3]	32.89%	50
[4]	25.66%	39
[5]	9.21%	14
TOTAL		152

BASIC STATISTICS				
Minimum	Maximum	Median	Mean	Standard Deviation
1.00	5.00	3.00	3.01	1.12

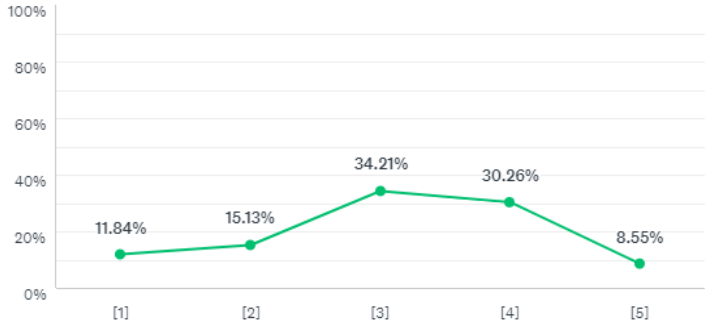
7.2. Collaborative Planning

ANSWER CHOICES	RESPONSES	
[1]	6.58%	10
[2]	21.05%	32
[3]	26.97%	41
[4]	32.89%	50
[5]	12.50%	19
TOTAL		152

BASIC STATISTICS				
Minimum	Maximum	Median	Mean	Standard Deviation
1.00	5.00	3.00	3.24	1.12

8.1. Total Quality Management (TQM)

Answered: 152 Skipped: 0

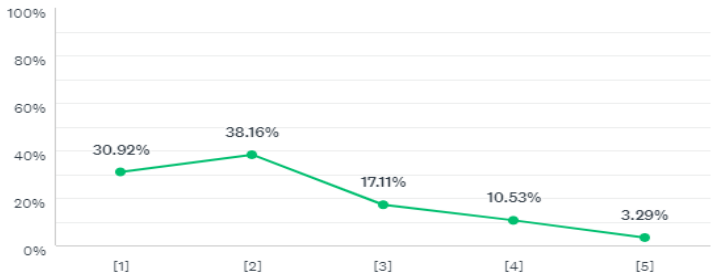


ANSWER CHOICES	RESPONSES	
[1]	11.84%	18
[2]	15.13%	23
[3]	34.21%	52
[4]	30.26%	46
[5]	8.55%	13
TOTAL		152

BASIC STATISTICS

Minimum	Maximum	Median	Mean	Standard Deviation
1.00	5.00	3.00	3.09	1.12

8.2. Error-Proofing (Poka-Yoke)

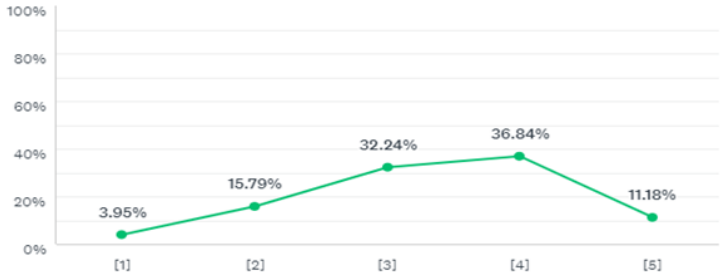


ANSWER CHOICES	RESPONSES	
[1]	30.92%	47
[2]	38.16%	58
[3]	17.11%	26
[4]	10.53%	16
[5]	3.29%	5
TOTAL		152

BASIC STATISTICS

Minimum	Maximum	Median	Mean	Standard Deviation
1.00	5.00	2.00	2.17	1.08

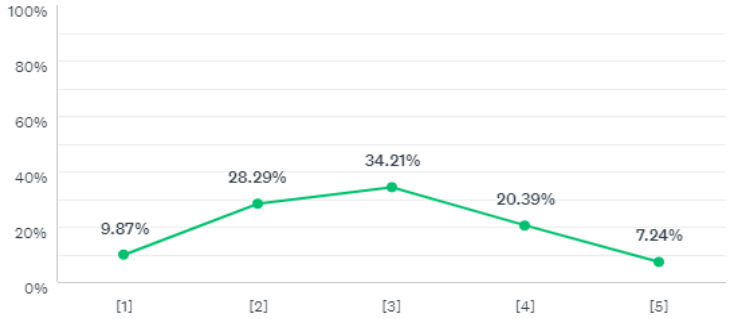
8.3. The Five Whys



ANSWER CHOICES	RESPONSES	
[1]	3.95%	6
[2]	15.79%	24
[3]	32.24%	49
[4]	36.84%	56
[5]	11.18%	17
TOTAL		152

BASIC STATISTICS				
Minimum	Maximum	Median	Mean	Standard Deviation
1.00	5.00	3.00	3.36	1.00

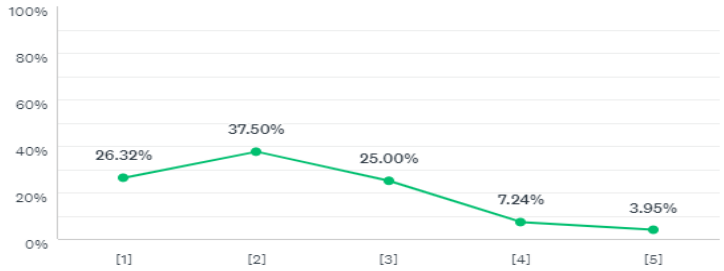
9.1. Work Organization



ANSWER CHOICES	RESPONSES	
[1]	9.87%	15
[2]	28.29%	43
[3]	34.21%	52
[4]	20.39%	31
[5]	7.24%	11
TOTAL		152

BASIC STATISTICS				
Minimum	Maximum	Median	Mean	Standard Deviation
1.00	5.00	3.00	2.87	1.07

9.2. Building Information Modeling (BIM)



ANSWER CHOICES	RESPONSES	
[1]	26.32%	40
[2]	37.50%	57
[3]	25.00%	38
[4]	7.24%	11
[5]	3.95%	6
TOTAL		152

BASIC STATISTICS

Minimum	Maximum	Median	Mean	Standard Deviation
1.00	5.00	2.00	2.25	1.05

Appendix B

Appendix B1: Computer Output from some of the *R* Procedures Used in this Study

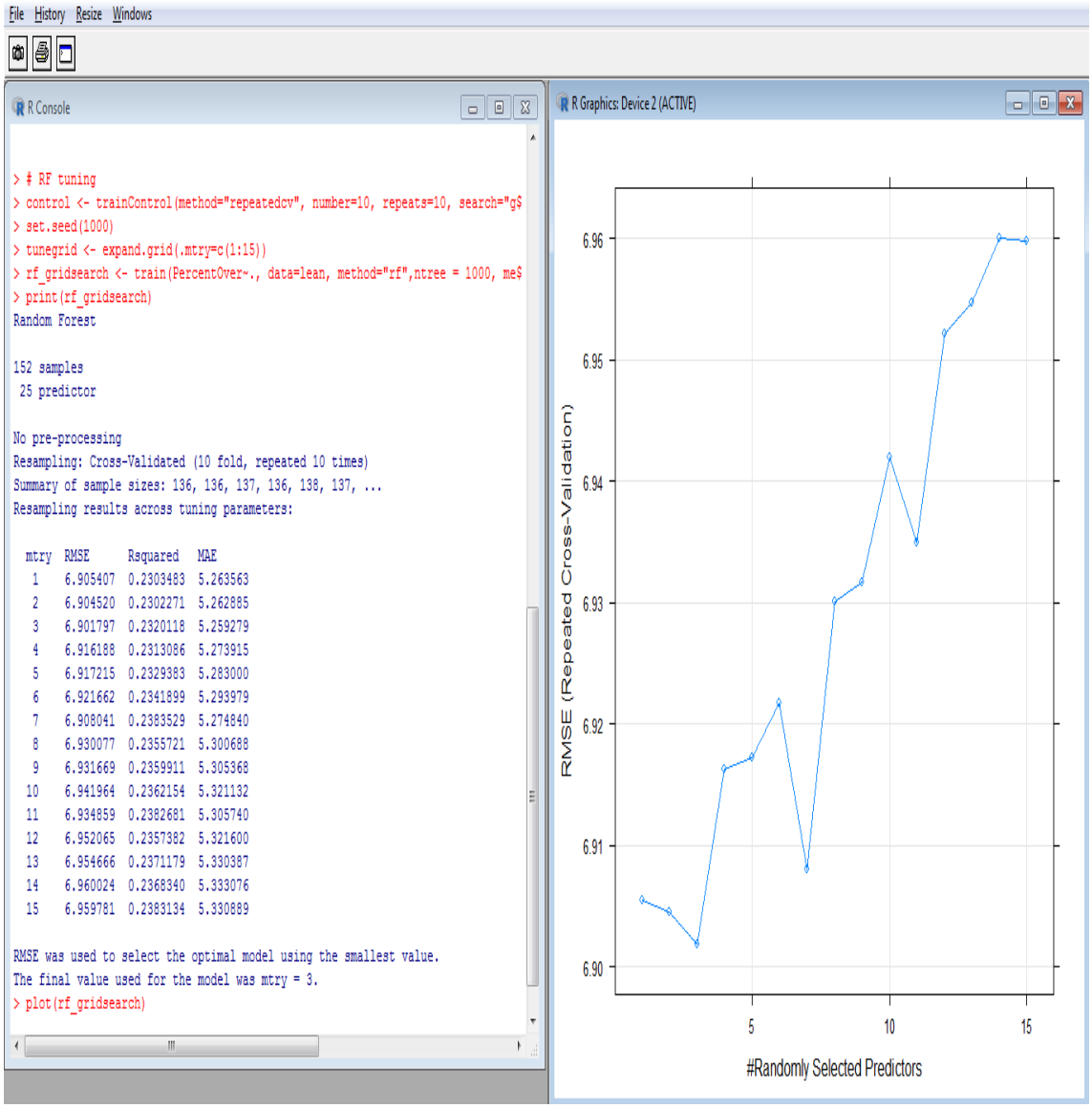


Figure 2: Predictors Randomly Selected by Applying Random Forest

Appendix B2: Structure of the Random Forest Model as an *R* Object

```
> str(RandomForestModel)

List of 18

 $ call           : language randomForest(formula =
PercentOver ~ ., data = lean, ntree = 1000, mtry = 3)

 $ type           : chr "regression"

 $ predicted      : Named num [1:152] 7.59 -2.02 7.89 7.07
7.94 ...

  ..- attr(*, "names")= chr [1:152] "1" "2" "3" "4" ...

 $ mse            : num [1:1000] 95.5 87.5 77.8 65.5 59.1
...

 $ rsq            : num [1:1000] -0.5974 -0.4631 -0.3004 -
0.0958 0.0116 ...

 $ oob.times      : int [1:152] 391 364 378 384 365 374 361
370 328 355 ...

 $ importance     : num [1:25, 1] 381 281 213 266 233 ...

  ..- attr(*, "dimnames")=List of 2

  .. ..$ : chr [1:25] "sp1.1" "sp1.2" "sp1.3" "sp1.4" ...

  .. ..$ : chr "IncNodePurity"

 $ importanceSD   : NULL

 $ localImportance: NULL

 $ proximity      : NULL

 $ ntree          : num 1000

 $ mtry           : num 3

 $ forest         :List of 11

  ..$ ndbigtree    : int [1:1000] 97 101 97 107 99 103 101
103 97 105 ...

  ..$ nodestatus   : int [1:117, 1:1000] -3 -3 -3 -3 -3 -3
-3 -3 -3 -3 ...

  ..$ leftDaughter : int [1:117, 1:1000] 2 4 6 8 10 12 14
16 18 20 ...

  ..$ rightDaughter: int [1:117, 1:1000] 3 5 7 9 11 13 15
17 19 21 ...
```

```
  ..$ nodepred     : num [1:117, 1:1000] 6.44 9.65 1.1 11.6
8.03 ...
  ..$ bestvar      : int [1:117, 1:1000] 25 1 25 25 11 18 9
5 24 22 ...
  ..$ xbestsplit   : num [1:117, 1:1000] 2.5 2.5 3.5 1.5
4.5 2.5 4.5 3.5 2.5 2.5 ...
  ..$ ncat         : Named int [1:25] 1 1 1 1 1 1 1 1 1 1
...
  .. ..- attr(*, "names")= chr [1:25] "sp1.1" "sp1.2"
"sp1.3" "sp1.4" ...
  ..$ nrnodes      : int 117
  ..$ ntree        : num 1000
  ..$ xlevels      :List of 25
  .. ..$ sp1.1: num 0
  .. ..$ sp1.2: num 0
  .. ..$ sp1.3: num 0
  .. ..$ sp1.4: num 0
  .. ..$ sp2.1: num 0
  .. ..$ sp2.2: num 0
  .. ..$ sp3.1: num 0
  .. ..$ sp3.2: num 0
  .. ..$ sp4.1: num 0
  .. ..$ sp4.2: num 0
  .. ..$ sp4.3: num 0
  .. ..$ sp5.1: num 0
  .. ..$ sp5.2: num 0
  .. ..$ sp5.3: num 0
  .. ..$ sp5.4: num 0
  .. ..$ sp6.1: num 0
  .. ..$ sp6.2: num 0
  .. ..$ sp6.3: num 0
  .. ..$ sp7.1: num 0
```

```
  .. ..$ sp7.2: num 0

  .. ..$ sp8.1: num 0

  .. ..$ sp8.2: num 0

  .. ..$ sp8.3: num 0

  .. ..$ sp9.1: num 0

  .. ..$ sp9.2: num 0

 $ coefs          : NULL

 $ y              : Named num [1:152] 12 -2 7 8 21 17 8 22
18 14 ...

  ..- attr(*, "names")= chr [1:152] "1" "2" "3" "4" ...

 $ test           : NULL

 $ inbag          : NULL

 $ terms          :Classes 'terms', 'formula'  language
PercentOver ~ sp1.1 + sp1.2 + sp1.3 + sp1.4 + sp2.1 + sp2.2
+ sp3.1 + sp3.2 +      sp4.1 + sp4.2 + sp4.3 + sp5.1 |
__truncated__ ...

  .. ..- attr(*, "variables")= language list(PercentOver,
sp1.1, sp1.2, sp1.3, sp1.4, sp2.1, sp2.2, sp3.1, sp3.2,
sp4.1, sp4.2, sp4.3, sp5.1, sp5.2,| __truncated__ ...

  .. ..- attr(*, "factors")= int [1:26, 1:25] 0 1 0 0 0 0 0
0 0 0 ...

  .. .. ..- attr(*, "dimnames")=List of 2

  .. .. .. ..$ : chr [1:26] "PercentOver" "sp1.1" "sp1.2"
"sp1.3" ...

  .. .. .. ..$ : chr [1:25] "sp1.1" "sp1.2" "sp1.3" "sp1.4"
...

  .. ..- attr(*, "term.labels")= chr [1:25] "sp1.1" "sp1.2"
"sp1.3" "sp1.4" ...

  .. ..- attr(*, "order")= int [1:25] 1 1 1 1 1 1 1 1 1 1
...

  .. ..- attr(*, "intercept")= num 0

  .. ..- attr(*, "response")= int 1

  .. ..- attr(*, ".Environment")=<environment: R_GlobalEnv>
```

```
  .. ..- attr(*, "predvars")= language list(PercentOver,
sp1.1, sp1.2, sp1.3, sp1.4, sp2.1, sp2.2, sp3.1, sp3.2,
sp4.1, sp4.2, sp4.3, sp5.1, sp5.2,| __truncated__ ...

  .. ..- attr(*, "dataClasses")= Named chr [1:26] "numeric"
"numeric" "numeric" "numeric" "numeric" ...

  .. .. ..- attr(*, "names")= chr [1:26] "PercentOver"
"sp1.1" "sp1.2" "sp1.3" ...

  - attr(*, "class")= chr [1:2] "randomForest.formula"
    "randomForest"
```

Appendix B3: *R* output for Variable Importance Matrix

```
> print (RandomForestModel$importance)

%IncMSE

sp1.1 8.188821

sp1.2 3.410581

sp1.3 3.985798

sp1.4 5.033972

sp2.1 7.452573

sp2.2 4.210911

sp3.1 8.022384

sp3.2 2.549738

sp4.1 4.939003

sp4.2 3.654249

sp4.3 1.180144

sp5.1 4.796970

sp5.2 16.187806

sp5.3 5.718841

sp5.4 6.195571
```

sp6.1 5.772828

sp6.2 5.986385

sp6.3 3.314713

sp7.1 7.121917

sp7.2 1.882002

sp8.1 4.024277

sp8.2 7.147191

sp8.3 6.826656

sp9.1 4.506955

sp9.2 7.935070

Appendix B4: Predicted Response for New Observations Using Model for Cost overruns

Coefficients

	Cost overrun	standardized cost
overrun		
Constant	15.7758	0.000000
Response3	-0.3473	-0.053656
Response4	-0.2590	-0.044418
Response5	-0.0056	-0.000865
Response6	-0.3323	-0.048835
Response7	0.5666	0.086249
Response8	0.0370	0.006234
Response9	0.6407	0.098561
Response10	0.0846	0.011767
Response11	0.3177	0.053174
Response12	0.3403	0.051173
Response13	-0.0141	-0.001988
Response14	0.1877	0.031228
Response15	-1.6203	-0.249840
Response16	-0.2335	-0.033047
Response17	-0.2441	-0.037545
Response18	0.4252	0.061691
Response19	-1.4455	-0.223739
Response20	-0.7495	-0.101979
Response21	-0.5238	-0.079966
Response22	0.0054	0.000821
Response23	0.1717	0.026732
Response24	-0.9026	-0.132839
Response25	0.1300	0.016930
Response26	0.1330	0.019287
Response27	-1.1953	-0.177092

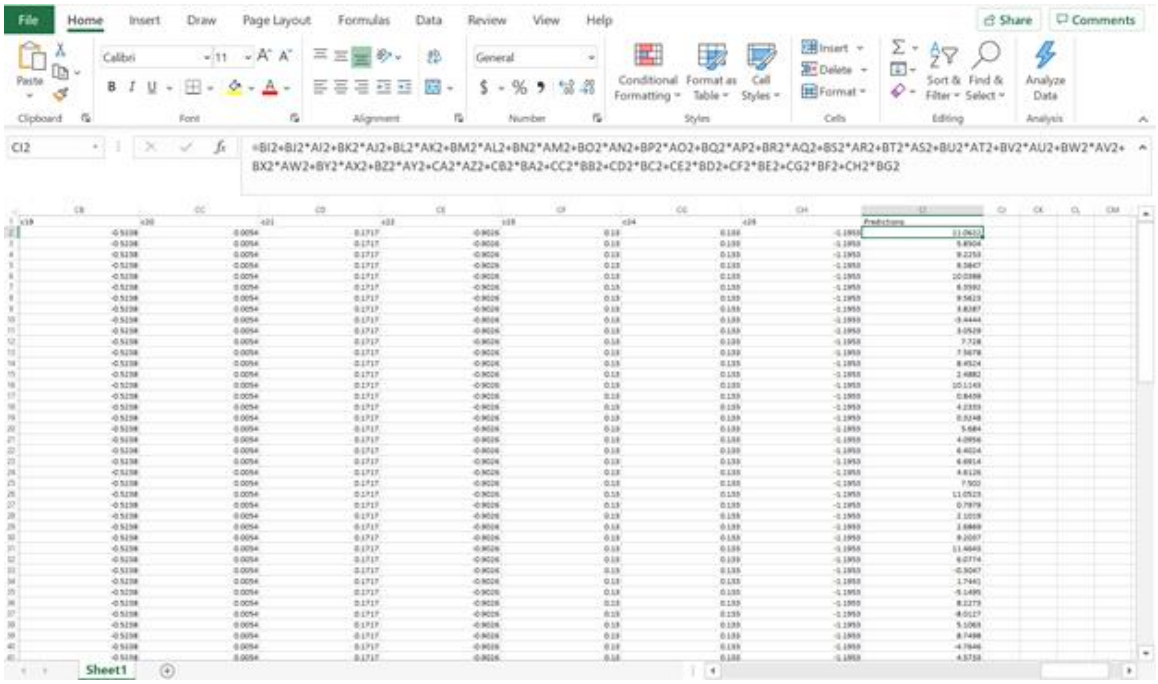
Predicted Response for New Observations Using Model for
Cost overrun

Row	Fit	SE Fit	95% CI	95% PI
1	11.0616	1.33004	(8.4228, 13.7003)	(-0.1752, 22.2983)
2	5.8498	0.55863	(4.7415, 6.9581)	(-5.1289, 16.8284)
3	9.2248	0.72197	(7.7924, 10.6572)	(-1.7912, 20.2409)
4	9.3841	0.95655	(7.4863, 11.2819)	(-1.7021, 20.4703)
5	10.0383	0.77993	(8.4909, 11.5856)	(-0.9933, 21.0699)
6	6.3585	0.57785	(5.2121, 7.5050)	(-4.6240, 17.3411)
7	9.5617	1.09682	(7.3856, 11.7377)	(-1.5755, 20.6989)
8	3.8282	0.60300	(2.6319, 5.0246)	(-7.1596, 14.8161)
9	-3.4453	1.17543	(-5.7773, -1.1133)	(-14.6140, 7.7234)
10	3.0524	0.58419	(1.8934, 4.2114)	(-7.9315, 14.0362)
11	7.7275	0.64999	(6.4379, 9.0171)	(-3.2709, 18.7259)
12	7.5672	0.89797	(5.7856, 9.3487)	(-3.4997, 18.6340)
13	8.4519	0.73783	(6.9881, 9.9158)	(-2.5683, 19.4721)
14	2.4878	0.69880	(1.1014, 3.8742)	(-8.5224, 13.4979)
15	10.1138	0.94017	(8.2485, 11.9791)	(-0.9669, 21.1944)
16	0.8432	1.06713	(-1.2740, 2.9603)	(-10.2826, 11.9690)
17	4.2330	1.01136	(2.2265, 6.2395)	(-6.8723, 15.3383)
18	0.3243	0.80047	(-1.2638, 1.9124)	(-10.7131, 11.3617)
19	5.6836	1.12875	(3.4442, 7.9230)	(-5.4661, 16.8333)
20	4.0949	0.62288	(2.8592, 5.3307)	(-6.8973, 15.0871)
21	6.4020	0.95540	(4.5066, 8.2975)	(-4.6837, 17.4878)
22	6.6910	0.71128	(5.2799, 8.1022)	(-4.3223, 17.7043)
23	4.6121	0.69514	(3.2330, 5.9912)	(-6.3972, 15.6213)
24	7.5015	0.82500	(5.8648, 9.1383)	(-3.5430, 18.5460)
25	11.0519	0.92511	(9.2166, 12.8873)	(-0.0237, 22.1276)
26	0.7972	0.77101	(-0.7324, 2.3269)	(-10.2319, 11.8263)
27	2.1012	0.64049	(0.8305, 3.3719)	(-8.8950, 13.0974)
28	2.6864	0.65337	(1.3902, 3.9827)	(-8.3128, 13.6856)
29	9.2032	0.72127	(7.7722, 10.6341)	(-1.8127, 20.2190)
30	11.4641	1.01033	(9.4596, 13.4686)	(0.3592, 22.5690)
31	6.0770	0.78805	(4.5135, 7.6404)	(-4.9569, 17.1108)
32	-0.3054	0.80378	(-1.9001, 1.2893)	(-11.3437, 10.7329)
33	1.7433	0.74766	(0.2600, 3.2267)	(-9.2795, 12.7661)
34	-5.1503	1.28965	(-7.7089, -2.5916)	(-16.3685, 6.0680)
35	8.2268	0.65310	(6.9310, 9.5225)	(-2.7724, 19.2259)
36	-8.0136	1.54533	(-11.0795, -4.9477)	(-19.3583, 3.3310)
37	5.1059	0.87450	(3.3710, 6.8409)	(-5.9535, 16.1654)
38	8.7491	1.02706	(6.7114, 10.7868)	(-2.3619, 19.8601)
39	-4.7655	1.25531	(-7.2560, -2.2750)	(-15.9684, 6.4373)
40	4.5728	0.67434	(3.2350, 5.9107)	(-6.4313, 15.5770)
41	-1.0117	1.14625	(-3.2858, 1.2624)	(-12.1685, 10.1451)
42	-5.0669	1.31360	(-7.6730, -2.4607)	(-16.2960, 6.1623)
43	4.3594	0.78327	(2.8055, 5.9134)	(-6.6731, 15.3919)
44	2.8067	0.59815	(1.6200, 3.9934)	(-8.1801, 13.7935)
45	8.6815	1.44744	(5.8098, 11.5532)	(-2.6122, 19.9752)
46	7.2758	0.85940	(5.5708, 8.9808)	(-3.7790, 18.3306)
47	6.6121	0.91663	(4.7935, 8.4306)	(-4.4608, 17.6849)
48	4.3580	0.76987	(2.8306, 5.8854)	(-6.6708, 15.3868)

Appendix B5: Calculation of MSE and RMSE

Observed Cost overrun (Y)	Predicted cost overrun (Y_hat)	Y-Y_hat	(Y-Y_hat)^2
21	11.0622	9.9378	98.75987
22	5.8504	16.1496	260.8096
14	9.2253	4.7747	22.79776
7	9.3847	-2.3847	5.686794
15	10.0388	4.9612	24.61351
13	6.3592	6.6408	44.10022
7	9.5623	-2.5623	6.565381
4	3.8287	0.1713	0.029344
0	-3.4444	3.4444	11.86389
0	3.0529	-3.0529	9.320198
5.2	7.728	-2.528	6.390784
2.4	7.5678	-5.1678	26.70616
6	8.4524	-2.4524	6.014266
4.6	2.4882	2.1118	4.459699
6.1	10.1143	-4.0143	16.1146
21	0.8439	20.1561	406.2684
3	4.2333	-1.2333	1.521029
-2.2	0.3248	-2.5248	6.374615
8.2	5.684	2.516	6.330256
-1	4.0956	-5.0956	25.96514
12	6.4024	5.5976	31.33313
1.3	6.6914	-5.3914	29.06719
2	4.6126	-2.6126	6.825679
11	7.502	3.498	12.236
3	11.0523	-8.0523	64.83954
12	0.7979	11.2021	125.487
0	2.1019	-2.1019	4.417984
-2	2.6869	-4.6869	21.96703
7	9.2037	-2.2037	4.856294
6.6	11.4643	-4.8643	23.66141
10.7	6.0774	4.6226	21.36843
19	-0.3047	19.3047	372.6714
-4	1.7441	-5.7441	32.99468
-0.1	-5.1495	5.0495	25.49745
18	8.2273	9.7727	95.50567
-1	-8.0127	7.0127	49.17796
12	5.1063	6.8937	47.5231
5.5	8.7498	-3.2498	10.5612
-1	-4.7646	3.7646	14.17221
8	4.5733	3.4267	11.74227

-4.5	-1.0109	-3.4891	12.17382
0	-5.066	5.066	25.66436
0	4.3601	-4.3601	19.01047
-2	2.8073	-4.8073	23.11013
5	8.6822	-3.6822	13.5586
9	7.2764	1.7236	2.970797
15	6.6127	8.3873	70.3468
40	4.3586	35.6414	1270.309

Table 3: Calculation of MSE and RMSE

Appendix B6: Calculating Predictions for the 30% Test Data from the Final PLS Model Equation