Gamification and Predictive Analytics for the Next Generation of Workers

By

Jalaja Shanmugalingam

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Engineering Management

Department of Mechanical Engineering

University of Alberta

Abstract

In this era of a new industrial revolution known as Industry 4, it is important for a company to determine ways to satisfy employees' needs and companies' needs simultaneously. This is the key to keep employees focused, engaged, motivated, and involved with their jobs. It would help companies stay competitive, as well as recruit and retain talent. The future of employee motivation is to take advantage of gamification elements using machine learning and a statistical model that have dominated social media and gaming applications. The implementation of psychological theories in the study of gamification has played an important role in our digital lives as multiple social media and gaming apps compete to harness user engagement to stay popular and relevant. A review of psychological theories that induce extrinsic and intrinsic motivations in the workplace, as well as the current applications of game elements in a non-gamified environment, are included in this research. Following that, a modified gamified model is developed in this research that uses real time data, Weibull statistical distribution, a K-means clustering algorithm, and machine learning along with a reward system that would help increase skill development and retention along with employee satisfaction.

Preface

This thesis is the original work of Jalaja Shanmugalingam under the supervision of Dr. Yongsheng Ma. The general direction of the research was suggested by Dr. Yongsheng Ma; while the specific topics, framework, and methodologies were developed by Jalaja Shanmugalingam.

Acknowledgments

Foremost, I would like to express my deepest gratitude to my supervisor, Dr. Yongsheng Ma, for believing in me and giving me an opportunity to take on this journey when I first reached out to him. I am thankful for his support, motivation, guidance, and immense knowledge throughout my master's study.

Secondly, my sincere thanks to Mr. David Lario and Enigma Design Solutions where this research idea first initiated. They generously allowed me to use their in-house software to gather the real-time data upon which the developed models are built. This dream of doing my master's survived early financial strains with the help of the initial funding obtained through the MITACS Accelerate Grant sponsored by Enigma Design Solutions.

I would also like to thank the Department of Mechanical Engineering staff at the University of Alberta and all the other professors with whom I have taken some of the required courses. The wealth of knowledge they shared during their lectures has helped me through my research. In addition, I am also grateful for the support and the encouragement I received from my colleagues and my supervisors at GPRC (Grande Prairie Regional College) in the past two years. They all inspired me to continue regardless of any hardships I faced.

Finally, my sincere gratitude to my parents, brothers, nephews, niece, and my friends for all the support and love. I would like to dedicate this thesis to my mother, who stayed up nights every week when I made the trip to Edmonton to make sure I arrived home safe regardless of how late it was, and always made sure to send me back to Edmonton from Grande Prairie with a week's worth of home cooked meals. Ever since I was a little girl, both my parents instilled in me the importance of education and taught me to follow my heart. My mother, especially, supported me in pursuing my dreams in every little way she could, to make sure I did not face the same hurdles she did when she was a young girl. For that, I am forever grateful.

Table of Contents

List of Figures

List of Tables

1 Introduction

Businesses started exploring the possibility of using game elements outside of customary context as early as the 1950s, soon after World War II ended, to optimize various human behaviours such as productivity, decision-making and cooperation. In modern days, the implementation of playful game elements in non-game-based environments to improve user motivation and involvement is known as gamification. In general, it is used by organizations to improve customer as well as employee engagement. This research focuses on the gamification of workplaces "… that use gamification to transform some of their work processes into a game-like experience for the employees by applying selected principles of game design and game interaction" [1]. This is done through the application of social and behavioural theories with the use of computer technology. Gamification has evolved over time in its methodology, purpose, and type of participants. Studying this evolution is imperative to understanding the needs of industry in developing an improved model.

Social scientists spent years researching human behaviour and human motivation, and developed theories to explain their findings. Gamification is one such example where psychology and technology came together to develop a system to improve human motivation, productivity and work experience. The future of gamification looks bright, as computer scientists are still finding new and improved ways to fully implement the existing scientific theories as the

technology grows. In this research we explored the diverse types of gaming simulations implemented throughout the last few decades and developed a gamification methodology that uses big data, statistical analysis, and machine learning to implement game elements in non-gaming contexts. The tracker used for data collection was developed and provided by Enigma Design Solutions. The data assortment, data analysis and model development using Python, MySQL,VBA and SQLAlchemy were contributed solely by the author. How this model uniquely differs from the existing models is discussed in detail in the chapters below.

In order to implement these theories using big data, one needs to collect behavioural data from users, which then can be analyzed using a rules engine. With this collection of data comes the great responsibility of getting proper consent to collect and protect personal data. There were no existing consent laws in place when this new phenomenon of online data collection hit the market. However, governments have slowly started legislating on this matter, especially after the recent revelations involving the collection and use of personal data through social media platform Facebook [2]. Therefore, gamified system developers must pay special attention to those laws that govern the collection of personal data in the respective territories in which they intend to use the gamified model.

2 Literature review

The concept of applying game elements in non-gaming contexts is not a new phenomenon.

This section explores the evolution of gamification by studying how the methodology, purpose, and type of participants changed along with the modern application of gamification in today's industry. This literature review will also include an in-depth exploration of current gamification methodologies in terms of their theoretical backgrounds and technological perspectives.

2.1 Theoretical Background

As a famous saying from an unknown author goes, "Choose a job you love, and you will never have to work a day in your life." It is important to choose a job one loves; however, passion for one's work cannot guarantee constant high levels of performance, energy, and motivation at work. This is where gamification comes into play. It is intended to enhance the playfulness of work to help motivate workers. Studies show that the number of gamers around the world continues to grow at a rapid rate and players, especially from the younger generation, are spending more time and money on games than ever before. The chart below shows the current and past global revenue from the gaming industry and the forecasted revenue up to the year 2020 in billions [3].

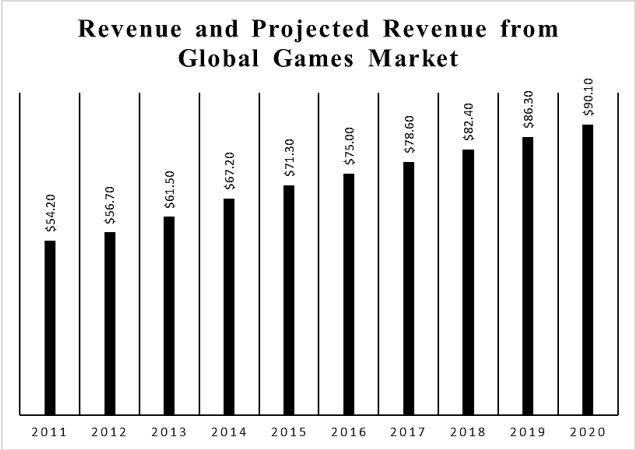

Figure 1: Revenue and Expected Revenue from the Global Games Market [3]

Even when the global economy struggled, the gaming industry saw a continued growth throughout 2012 to 2019 [3]. What makes these players continue to voluntarily spend their valuable resources: money and time? Psychologists studied gamers' behaviour and developed the following theories to generalize the factors that induce motivation in gameplay:

- Users are considered to be motivated when they are engaged, focused, and involved in what they do [4].
- A motivated user tries harder to achieve goals, and his/her choice to excel is self-driven and strong-willed [4].
- Motivation can arise from both internal and external factors.
- The social aspect is also considered an influential motivational factor.

2.1.1 Intrinsic Motivation

Intrinsic motivation is when a user is motivated through internal factors. For example, doing something that would help an employee feel like an asset to the company could be considered an intrinsically motivated action. This type of motivation is explained using several needs-based theories, such as Maslow's hierarchy of needs, need achievement theory from Atkinson, goal setting theory, and self-efficacy theory by Bandura [5].

Maslow's hierarchy of needs, published in 1943, argues that a certain level of needs must be fulfilled before one is motivated to achieve the next level of needs [6]. For example, the first level of needs in the hierarchy is physiological needs, such as breathing, eating, and getting enough sleep. Once these needs are fulfilled, then the person would go onto the next level. The complete hierarchy is shown below [6].

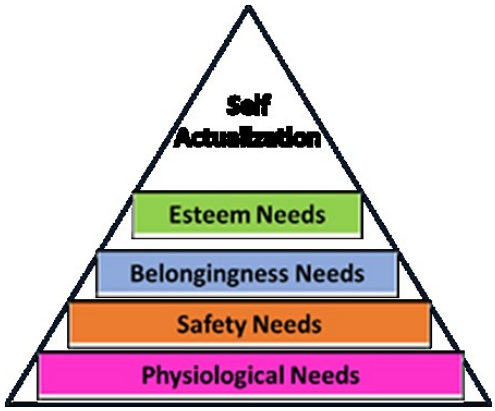

Figure 2: Hierarchy of Needs [6]

Studying these needs, Siang et al. developed the game players' hierarchy of needs as follows[7].

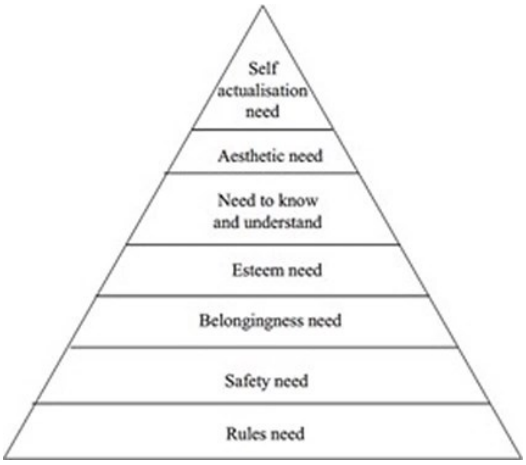

Figure 3: Game Players' Hierarchy of Needs [7]

In need achievement theory, the authors describe the achievement behaviours as building or showing high ability [8]. As a result, people would prefer success as opposed to failure; that is, choosing tasks that would guarantee success. If this were to be the motivation, then stronger people would choose simpler tasks or tasks with highest difficulty [9]. To make use of these motivational factors, games usually have built-in achievement systems that help players identify their goals and achievements along with the rewards for every achievement. As the player achieves his/her goals and gains rewards, he/she is motivated to go further and do more. Some games also allow the player to choose the difficulty for set tasks. Once they start getting comfortable with the rules of play and see the fruit of their labour, players do choose to seek a greater challenge that would give them better internal rewards. Goal setting theory is based on a similar principle.

Goal setting theory states that immediate goals that are hard, specific, and situation-

appropriate would motivate a player to strive to achieve more long-term goals [10]. The theory of self-efficacy is based on the same idea. According to Bandura, self-efficacy is one's supposed ability to perform a task. Self-efficacy, depending on how strong it is, can either improve or hinder motivation [5]. Just like in goal setting theory, "…self-efficacy can be positively stimulated by dividing tasks of higher difficulty into smaller, less difficult tasks…," because it would enhance performance experience; self-efficacy is thus considered one of the most influential motivational factors [11].

2.1.2 Extrinsic Motivation

Extrinsic motivation, on the other hand, is the desire to do something in order to achieve external rewards such as badges, points or even bonuses in some cases. This type of motivation is characterized by expectancy-value theory (EVT), and Skinner's principle of partial reinforcement. According to EVT," …Goal-directed behavior is a function of the belief that efforts will lead to performance needed to attain the rewards" [12]. Skinner's principle of partial reinforcement states, "…Continuous reinforcement establishes desired behaviors quicker than partial reinforcement" [12].

Finally, a comprehensive theory called self-determination theory focuses on the type of motivation instead of the quantity. It focuses on three different psychological needs: "…autonomy, competence, and related-ness…" [12]. Autonomy is an extrinsic motivation factor, whereas competence is an intrinsic motivation factor and related-ness is the social aspect

that the theory focuses on. Current gamification models have all three components integrated into their systems.

2.1.3 Social Aspect

In addition to the above-mentioned theories and intrinsic/extrinsic motivational factors, the social aspect of gamification is also an important motivational factor. Social theories such as social comparison and personal investment theory explain the motivations that arise from social factors and the impact these factors have on players. “Social comparison states that people seek to evaluate their beliefs, attitudes and abilities by comparing their reaction with others. Personal investment theory suggests that the level to which a person will invest personal resources of effort and time for an activity depends on personal incentives, beliefs regarding oneself, and comprehended alternatives” [12].

2.2 Evolution of Methodology

Gamification is, according to [13], “A process of enhancing a service with affordances for gameful experiences in order to support user's overall value creation.” The meaning of gamification has not changed much over the years; however, the application has evolved over time, as technological advancements have allowed companies to incorporate numerous conditions into their methodologies. This part of the literature review section strictly looks at the structure of gamification, such as number of players, rules of play, and use of technology; and focuses on how all these aspects have changed over time.

2.2.1 First Instances of Business Gaming

The first modern-day computerized business game, Top Management Decision Simulation, was introduced by the American Management Association (AMA) in 1956 to train upper level management [14]. In this mathematical game model, cause and effect formulas are used to predict the outcome of players' moves. In addition to the mathematical formula used, the following are some of the key components that define the simulation methodology of the game environment [14].

- No direct human interaction is involved except for role-playing.
- No humanics elements, such as motivational values or organizational problems, are included.
- Pre-defined goals are to be achieved in a sole-product industry for a familiar customer market where every firm involved has access to limited information on their competitors.
- The outcomes are purely management-oriented. Some of the output variables include production volume, product price, sales volume, advertising budget, position of inventory, Research & Innovation and Marketing budget.

AMA's Top Management Decision Simulation was also the first business game that used a computer, the IBM 650, to calculate participant score. This is believed to have helped remove the subjectivity that arose from human scoring and expedite the scoring calculations. Although

the scores were manually entered, thereby leading to errors in the outcome [15], this was still the first step to reduce the human subjectivity in evaluation, and laid out the path to modern-day gamification. Following the AMA's successful implementation of a business gaming simulation, numerous similar games came out. According to Faria et al., there existed more than 100 business games in the United States alone by 1961 [16]. From this point on, the number of business games grew rapidly in various industries along with technological development. The advancement in technology and invention of supercomputers helped develop business simulations with complicated mathematical models and dynamic variables. Thus, it can be concluded that technology played a very important role in shaping the game methodology. As described by Wolfe in his publication, the following are the distinct phases of business game development that came about due to technological advancements [17].

Phase I (1955 to 1963): Creation and growth of hand-scored games.

Phase II (1962 to 1968): Creation of mainframe business games and growth of commercially published games.

Phase III (1966 to 1985): Period of fastest growth of mainframe games and significant growth in business game complexity.

Phase IV (1984 to 2000): Growth of PC based games and development of decision-making aides to accompany business games.

Phase V (1998 to present): The growth of business game availability on the internet and run

through central servers (e.g., Capsim and the Capstone series of business games and Innovative Learning Solutions, as well as the Marketplace simulations).

The affordability of personal computers and the invention of the world-wide-web revolutionized the way business games were played. The Internet increased the accessibility of business games, thus increasing the number of participants. However, during the initial stages of web-based business games, there were concerns regarding the security of data collected because, "Prior to 2003,…, most web-based simulations were not yet fully online … data needed to be downloaded to local computers and then uploaded to the server program" [18]. However, recently developed web-based business games are run from a central server where the game decision variables are selected and simulated by the game administrator and the participant contributions are directly entered to the central server. In other words, high-speed internet and the advancement in server technology allows the collection and the processing of data from one central location, thereby increasing the data security [18].

Advanced technology also allowed game designers to model and incorporate more gaming that revolutionized the purpose of business games as an educational and training tool.

2.2.2 Purpose

The learning objectives of business games have changed over time. Most of the business games in the initial stages were management games aimed at training top-level executives. However, as virtual entertainment gaming became popular, business game developers began

incorporating those elements that captivate players to motivate and train all levels of workers. According to [16], who studied over 304 gaming articles, the Rank Order of the Five Major Education and Learning Objectives have changed over time, as shown in the table below.

Objectives	**1970s**	**1980s**	**1990s**	**2000s**
Learning Outcomes and Objectives	1	2	5	3
Decision-making skills	2	4	2	4
Teamwork	3	5	4	5
Experience gained	4	3	3	1
Strategy formulation	5	1	1	2

Table 1: Rank Order of the Five Major Education and Learning Objectives [16]

Ever since their inception, business games have been used as educational tools in a controlled environment where parameters were controlled by the game administrators. The application of business games in training people was the first use of computerized game elements in a work environment. This led to the modern-day reward-based gamification of workplaces where companies use rewards to harness certain positive behaviours from their employees [12].

2.2.3 Types of Participants

At the time when businesses started implementing game elements, it was to train the people from upper management to help them learn the cause and effect of their decisions. However, the

21st century business world changes so rapidly that even lower level frontline staff is sometimes faced with a situation of making a crucial decision that may have a significant impact on any business. Therefore, "Addressing these challenges requires a wide range of skills from both senior and front-line staff, in-turn requiring innovative and effective training tools such as serious games, gamification applications to aid staff at all levels of an organization as they adapt in response to emerging challenges" [19].

2.2.4 Modern Business Games

Modern business games place a huge emphasize on including game elements that would improve and enhance a wide range of skills, including soft skills such as communication, cooperation, and teamwork. According to Deci et al., one can be motivated through internal factors, such as a desire to be successful, or external factors like coercion, as discussed in the previous section [20]. The following are some of the areas where modern gamification is being applied [21]:

- Customer relations
- Improving employee performance
- Training
- Education
- Research and Innovation

- Personal development
- To encourage a healthy lifestyle

The behaviours one would want to gamify should be the ones that would optimize productivity. Productivity does not just depend on how much work is being done by whom, but also the quality of the work, as well as the extent to which goals set by employees to complete their tasks align with a company's short-term and long-term goals. Some examples of current gamification models are given in detail in **Section 2.3.4**.

As glamorous as gamification might look, there are still some shortcomings and ethical issues that should be addressed. According to Kim & Werbach, "Gamification can also unintentionally, recklessly, negligently, or inadvertently encourage players to cause harms to involved parties" [22]. Therefore, it is important for any companies that use gamification presently or design gamification for future to continuously evaluate the system and the actions of players to come up with a model that would help avoid these ethical issues.

2.3 Technological Overview of Gamification: Functional Requirements

This section aims to gain further knowledge of gamification elements that are being used from a technological perspective. Successful implementation of gamification depends on how well the researchers can develop a software model that would help implement the game elements in real-world scenarios. Over the years, psychologists have studied the reasons behind human motivation and provided various theories that closely describe why one may or may not be motivated to do

certain tasks, as discussed in the previous section. Now, computer scientists take these theories and develop software models to harness the power of human motivation. At this point, the quality of the software model is what determines whether the implementation of gamification will be successful or not. There is no commonly accepted taxonomy in gamification design as of yet. However, according to Deterding et al., gamification design is defined using five different levels as listed below [23].

1. Game Interface Design
2. Game Design Patterns
3. Game Design Principles
4. Game Design Methods
5. Game Models

2.3.1 Game Interface Design

This is the portal through which end users interact with a gaming system. This is one of the most crucial elements of any game development. Badges, levels, and leaderboards are some of the features that are included in game interface design in modern day gamification.

Figure 4: Example of Badges

Figure 5: Example of Levels [24]

Last Name	Student Name	Rank		Total XP	AdVENTURER	Blog Level	Geologic Time	Evolution	Charity Fair
Last name 1	Name 1		Trapper	15330					
Last name 2	Name 2		Apprentice	1291					
Last name 3	Name 3		Ranger	59350					
Last name 4	Name 4		Guardian	126994					
Last name 5	Name 5		Ultimate survivor	785610					
Last name 6	Name 6		Ranger	70222					

Figure 6: Example of Leaderboards [25]

2.3.2 Game Design Patterns

Design pattern, according to Christopher Alexander, "… describes a problem which occurs over and over again in our environment, and then describes the core of the solution to that problem in such a way that you can use this solution many times over, without ever doing it the same way twice" [26]. In his paper, Alexander generalized design patterns for towns, buildings, and constructions. However, this principle can still be applied to object-oriented programming; in this context, "design patterns" describes the objects and classes that are designed to solve a problem and the communication between them. According to Gamma, there are four elements to design patterns: patterns, problems, solutions, and consequences [27].

In cases of gamification of work, patterns represent when, where, why, and how an employee does a job [28]. An example of a problem could be employees being uninterested in completing certain tasks that are optional and repetitive, yet crucial to their jobs, such as staying informed on updated policies. The solution could be gamifying the learning process with badges and points for employees to collect every time they take and complete certain courses. This would keep them engaged as they obtain extrinsic rewards for their involvement.

2.3.3 Game Design Principles

Game design principles are "Evaluative guidelines to approach a design problem or analyze a given design solution" [29]. Some examples of design principles are "enduring play, clear goals, and variety of game styles" [30].

2.3.4 Game Design Methods

The creative, conscious solutions to solve existing problems using gamification are known as game design methods, which include very specific practices. One of the main reasons why companies look into gamifying certain tasks is to make the tasks more fun; this makes users feel more engaged and focused, inducing a mental state known as "flow." According to Csikszentmihalyi, flow is "A state in which people are so involved in an activity that nothing else seems to matter; the experience is so enjoyable that people will continue to do it even at great cost, for the sheer sake of doing it" [31]. Therefore, the role of a designer is to include design features that are fun. Flow state exists when the following three requirements are fulfilled [31]:

- Performing activities that are challenging
- Specific goals and instant feedback are offered
- Uncertain outcomes that can be influenced by players' actions ("paradox of control")

The following are the effects of flow state on a player [32]:

- A merging of action and awareness
- Concentration on immediate tasks
- Loss of awareness of self, loss of ego
- A distorted sense of time
- The experience of the activity is an end in itself; it is done for its own sake and not for an

external reward.

The following game design methodology can be used to implement the flow state and have a player engaged longer [32].

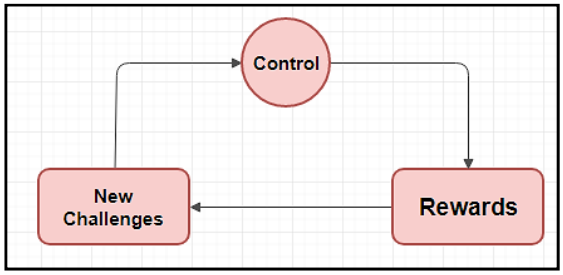

Figure 7: Game Design Flow

The following example of an eLearning course shows how the above model is applied in an actual gamified system [33]. “This gamified eLearning course featured scenario-based learning, where learners have to help their clients understand the features and functions of a software product. Each topic in the course was a level, which the learners have to progress through, completing challenges and earning experience points, badges, and trophies along the way. It also features dashboards and leaderboards to impart a sense of social recognition in learners” [33].

Figure 8: Gamified eLearning System [33]

The following are some additional popular gamified examples in the business world:

Company	Description of the system
Melbourne Storm: Loyalty Rewards Program	This is a loyalty program aimed at rewarding club members and fans who show their support for the team for a range of actions [34].
Engine Yard	Customer and Employee reward program for their contributions [35].
CallidusCloud: Mysalesgame	Designed to improve salesperson performance through training and adoptive behaviours [36].
Deloitte	"Deloitte gave their consultants instant feedback of their

	progress and guided them along clear learning paths. As consultants completed related coursework, they built up reputation that intrinsically motivated them to continuously engage" [37].
Bridge	This is another learning and training platform that offers a gamified system that includes the following components: online courses, in-person training, interactive videos, ongoing performance management, and assessment of skills [38].
Tovuti	Another eLearning software that helps organizations create eLearning programs to aid in learning and development. For this platform, the end user gives the information needed to build the training courses they need [39].
More examples of learning and development systems	SABA Cloud, Skyprep, Docebo, and Prosperity LMS [40] [41] [42][43].
Gameeffective	This platform is designed to help improve sales, learning and development for employees and customer service. It uses real-time data derived from other enterprise applications such as SAP. Using those data, it gives real time feedback to

	employees. Some organizations that use Gameeffective are as follows [44]: **Verizon:** To improve sales performance **Comedata:** To improve employee accountability and engagement **Microsoft:** To drive call center agents' performance.

Table 2: Examples of Gamified Systems

As can be seen above, most of these systems focus on external customers. At present, only very limited systems are available that focus on employee engagement and growth. Even those ones are mostly learning and development gamified systems.

2.3.5 Game Models

Game models are the conceptual models of gaming components such as challenge, fantasy, and curiosity. The figure below shows one of the well-researched, well thought out and complete gamification frameworks from Yu-Kai Chou according to this conceptual model. Chou's framework uses eight core drives of gamification [45]:

1. Epic Meaning & Calling
2. Development & Accomplishment
3. Empowerment of Creativity & Feedback
4. Ownership & Possession

5. Social Influence & Relatedness
6. Scarcity & Impatience
7. Unpredictability & Curiosity
8. Loss & Avoidance

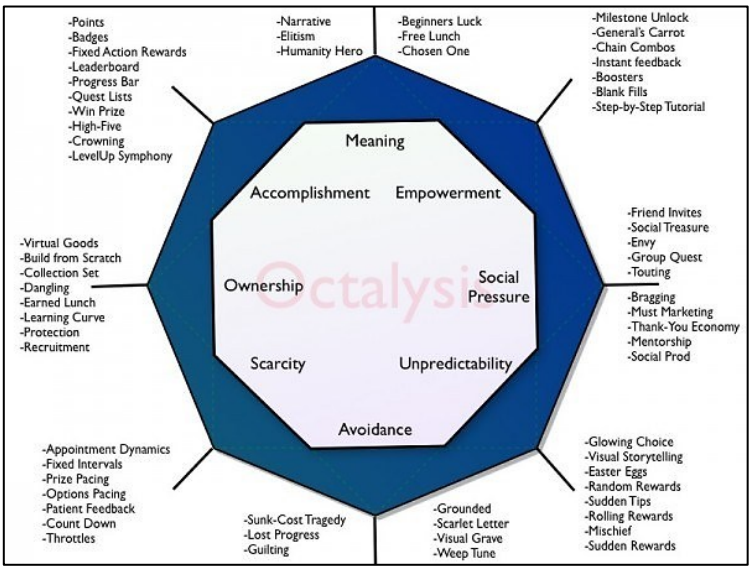

Figure 9: Octalysis [45]

3 Problem Statement

The current economy is living through the 4th Industrial Revolution that began in 2000, which is also known as Industry 4 [46]. Due to its innovative nature and rapid growth, companies are facing challenges in being able to adapt and grow. The challenges do not just stem from industrial requirements and customer expectations, but also from the generational changes that are happening in the workforce. This research focuses on finding solutions to some of the main

challenges faced by businesses and employees through:

- Developing an efficiency model using big data and a statistical model using Weibull distribution with real time data.
- Developing a gamified model that uses the output from the efficiency model to motivate, reward, and recognize employee contributions.

4 Development of Efficiency Model

4.1 Background study

Modern economies are divided into sectors, which are then classified into industries comprised of individual businesses grouped together based on their similarities. Examples of such classifications are shown below for the Professional, Scientific and Technical Services Sector, as well as some of the industries within that sector [47].

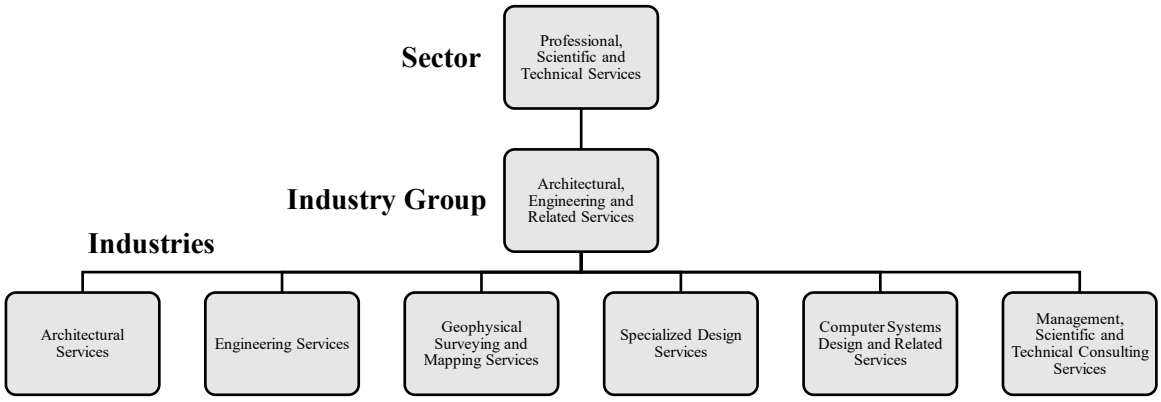

Figure 10: Professional, Scientific and Technical Services sector [47]

Every industrial revolution that has led to economic breakthrough and drastically improved human living conditions has stemmed from breakthroughs in a specific industry, as shown below [46].

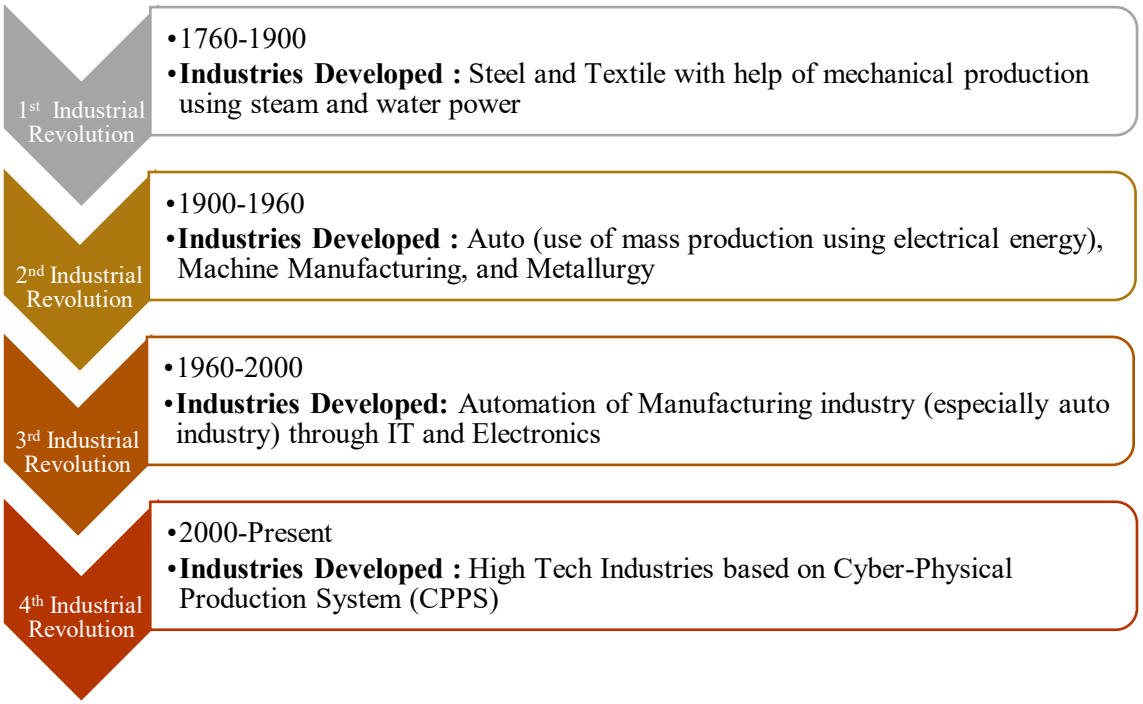

Figure 11: Breakdown of Industrial Revolutions [46]

The 4th Industrial Revolution is building upon the 3rd digital revolution in a more sophisticated way. At the turn of the 21st century, the development of machine learning and artificial intelligence allowed the easy integration of physical and virtual worlds. As a result, smart factories are becoming the center of new manufacturing, where quick product customization is made much more affordable through advanced technologies such as 3D printing technology and other additive manufacturing methods [48]. Unlike other industrial revolutions, the 4th Industrial

Revolution is happening in numerous industries; prominent ones include the advancement of nanotechnology in life sciences and breakthroughs in the use of renewables in the energy industry [4]. The simultaneous advancements happening in multiple domains make the current industrial revolution unique, and also demonstrate the magnitude of the revolution made possible through the advancements in digital technology. This revolution is progressing at an exponential rate compared to previous ones; businesses that have difficulty adapting to new changes could vanish from the market [49].

It is a well-known fact that the economies that were pioneers in the leading industries of each industrial revolution became powerful leaders of the world economy [50]. It is also evident that the countries that did not have the capacity to be part of every industrial revolution, due to systemic problems, are still having difficulties building strong economies. For example, 17% of the world population (1.3 billion people) has still not experienced the benefits of the 2nd Industrial Revolution and is living with no access to electricity [49]. Therefore, it is very important for a country's economy to understand the changing dynamics in various industries and spend enough resources on developing those industries that are leading the way to the future. In this fast-paced economic environment, agility and adaptability are what help a business stay relevant [51]. For that, a company should understand the market requirements and the challenges involved in meeting those requirements. In addition, companies should also understand the generational changes that are happening in different industries to be able to adapt to the needs

and habits of the new generation of workers. This would help them attract and retain human talent [52].

An industry is comprised of many individual companies and workers. The survival of industry, companies, and workers are interdependent. Hence, the challenges faced by different industries to stay relevant are not all the same; understanding and facing these challenges would help an economy continue to grow strong and remain steady even in times of global economic fluctuation.

4.1.1 Industry 4 and Challenges Faced by the Industries Leading the New Revolution

High tech industries based on Cyber Physical Production Systems are leading the 4th Industrial Revolution and radically changing the way things are done in every industry [53]. For a country to stay competitive in this globalized economy, it must make sure there are enough resources available for those industries that are backbone of their economy to function and grow, especially human capital [48]. Otherwise, those industries would die out and businesses that fuel them would leave in favor of new locations with affordable resources. There are various reasons why a business can struggle economically; they stem from social, political, and economic systems. One of the main reasons is the difficulty faced in acquiring qualified human capital to push forth growth and innovation [54].

4.1.2 Challenges Faced by Organizations

Organizations must radically transform themselves to be able to take advantage of the new

digital revolution and be competitive in the market. They should not only integrate the inventions of digital revolutions in organizational management, but also be innovative in product development, service delivery and manufacturing [53]. The following are some of the challenges that businesses are facing as they are living through the 4th Industrial Revolution.

Widening skills gap: The baby boomers who have been running these industries for the past few decades are approaching retirement. Let's consider North America as an example; according to a study done by Deloitte, "The skills gap may leave an estimated 2.4 million positions unfilled between 2018 and 2028" [55]. Another study shows that manufacturing investments also help create jobs in other industries such as goods and services; to be exact, every manufacturing job generates 2.5 new jobs in industries that support production [55].

Increased demand for customization: As mentioned in the previous section, product customization is made easier through advancements in digital technology. In addition, there is an increasing expectation and demand for customized products among customers. Therefore, companies should have the ability to design and produce customized products to continue to stay competitive and relevant. For example, "Automobile makers such as Ford and Toyota offer friendly interfaces through which buyers can 'design' their own cars" [56].

Leadership challenges: Except for several tech giants such as Google, Apple, and Amazon, few leaders in the business world understand the new generation of workers and their expectations. As baby boomers are nearing their retirement age, they are being replaced by

millennials: a generation of workers who grew up with advanced technology and a more connected global environment. Trying to lead or manage them in traditional ways does not produce the expected results. Rather, this would lead to a loss of talent, as the new generation of workers is notorious for leaving or switching jobs if they feel their expectations are not being met [57]. In addition to that, organizations are also challenged with scaling up their trainings of new employees with limited number of mentors due to baby boomer retirement [58].

Increased social responsibility: In this era of social media, companies are expected to be more socially and environmentally responsible instead of just focusing on profits. In recent years, companies have undertaken multiple projects to show their societal involvement, such as female empowerment, youth empowerment, and ending poverty. This has now become an important strategy for many businesses in order to build relationships with customers and retain qualified employees who prefer working for socially responsible organizations [59].

4.1.3 Challenges Faced by Employees

As the new generation of workers takes over the workforce, it is important for companies to understand the challenges they face to help them navigate the complex economic environment. Following are some of those challenges, as identified by researchers:

Lack of mentorship: A Deloitte survey from 2018 states that 43% of workers are planning on leaving their jobs due to lack of mentorship [60]. Another survey done by the Huffington Post shows that 79% of millennial workers see mentorship as an important component of career

success [61]. Mentorship in itself is a difficult process. The mentor is removing themselves from their productive time to coach someone who will be competing for the same workload and advancement in the company. Moreover, most companies do not compensate, reward, or acknowledge these contributions.

Low motivation: Millennials are notorious for hopping and ghosting jobs. Some of the reasons for this are as follows: disbelief in conventional hierarchy, desire to avoid bureaucratic complications, and not getting enough credit for their involvement in an organization's advancement [62].

Difficulty with work-life balance: Work-life balance is an important contributor to a healthy and happy life. It helps decrease stress and increase productivity, focus, and creativity. For a company to attract and retain top talent, it should have a system in place that is flexible enough to offer an opportunity for this new generation of workers to get work done, rather than compelling them to work the traditional 8 hours a day, 5 days a week schedule [62]. By creating a system to support work-life balance, companies can reduce their financial losses by increasing productivity and decreasing the need for healthcare spending. For example, according to Harvard Business Review, "The psychological and physical problems of burned-out employees, which cost an estimated $125 billion to $190 billion a year in healthcare spending in the U.S., are just the most obvious impacts" [63].

4.1.4 Research Question

It is essential for a business to understand and adapt to generational changes and remodel their organizations in order to tackle the above-mentioned challenges. A 20th century solution to a 21st century problem will not be helpful. In some industries, the solution companies came up with to deal with labour and skill shortages was outsourcing, "…'outside resourcing,' meaning to get resources from the outside" [64]. However, is this a long-term solution to these problems? This might increase the market capital and generate wealth in the short-term, but it would not ensure the long-term survival and growth of an organization. If companies are not ready to invest in human capital locally, then skill development will not happen. The skill shortage would become an even bigger problem, because "The most important factor of economic and social development at different times is the consideration of the effective role of human resources, because unlike other creatures, humans possess creative and efficient role and it can improve the quantitative and qualitative aspects of their work and eliminates potential problems with new methods" [65].

The companies that stood the test of time and survived for decades are the ones that focused on long-term solutions to the challenges they faced, as opposed to concentrating their efforts on short-term financial gains. This paper explores a methodology of using big data to find a solution that would help solve numerous problems discussed above.

4.2 Model for Work Productivity

The main goal of the current research is to model an effective method to use quantifiable data

to optimize employee efficiency and the work environment. First, the optimization process of a work environment must be based on the set goals and parameters that define success or failure. Second, there must be a mechanism to gather data and interpret the results. With growing improvements in artificial intelligence and machine learning, computers can be used to sort through many records and interpret results that may not be apparent to manual evaluation.

The diagram below shows the process of such model development. As shown in the diagram, the model developed is a dynamic model, which constantly collects data and continues to shape its parameters based on the data collected. A detailed explanation follows in the sections below.

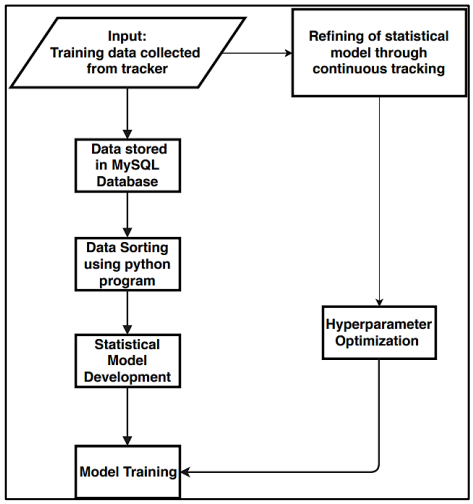

Figure 12: Process of Model Development

The three core components of the selected model are:

- Data Collection
- Data Analysis

- Decision Making

4.2.1 Data Collection

The main independent variable considered here is time. Among many other variables, time is the key factor that affects the success of businesses as well as employees. Effective time management helps with:

- Resource management (other than that of time)
- Decision making
- Productivity
- Setting measurable goals
- Learning opportunities
- Stress reduction
- Work-life balance
- Reduced financial losses
- Work quality
- Customer satisfaction

4.2.2 Data sorting and model development

Data collected for this specific research are based on time spent completing engineering

projects and the different components of that project. For example, let's consider projects as a universal set and an engineering design as one of the subsets with the following elements: modeling, analysis, calculations, and sometimes research, as shown in the Venn diagram below.

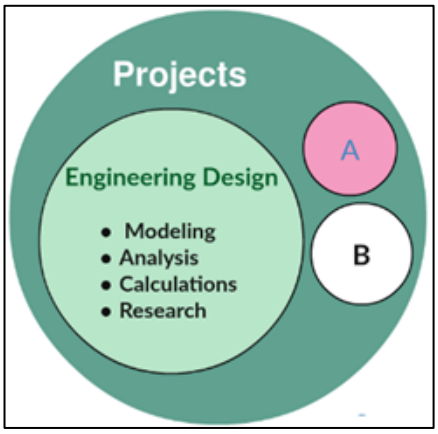

Figure 13: Projects and Subsets

In a typical project, if we consider a single task in detail, such as the solid modelling of a single component, each part created will have different features and each person creating the part may have a different process than others working on the same task. It has been historically argued that no useful patterns or meaningful information could be obtained from studying this drafting process. However, using data mining and artificial intelligence, complex patterns and habits can be extracted out of data that would not otherwise be apparent.

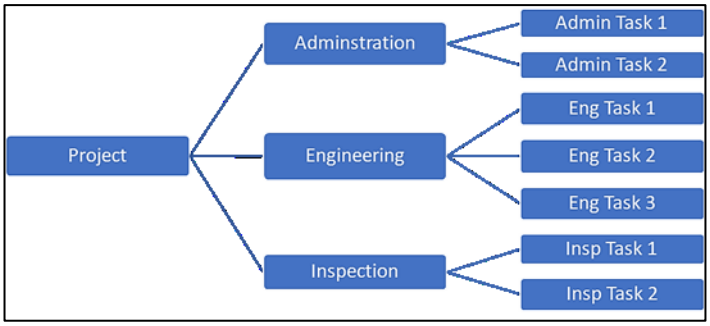

Figure 14: Components of an Example Project

Let's consider solid modelling of a simple component in Solidworks. For this component, the geometry is made by creating sketches and features from the sketches. With each feature created, new features can be built on top of them. Using CAD software based on parametric modelling, each step is stored in a hierarchical format with a strong parent-child relationship with sketches and features that depend on their parent. Any changes to the parent would have direct implications on the child features (i.e. changing geometry or removing parent). The other feature of parametric modelling is that each sketch or feature is an object with properties that can be extracted to determine things like number of types and quantities of a feature. It would be ideal to understand how long it takes to create each feature. This could be accomplished by directly measuring the editing time of each feature; if this is not possible, the editing time of the entire component could be measured, and a model could be developed on the assumptions of the procedures that were performed for creating the part.

The total time (T) for creating the item would be a summation of the time to create each feature (t_n).

$$\sum_{n}^{m} t_n = T \qquad (1)$$

If the time for each feature cannot be measured, and the total time taken to create a model is known, an assumption of "time to create" the feature could be used to find the time estimate for a specific feature completion. In this estimation, T is the total time, and A_n is the time-weighted average associated with it, calculated as shown below:

$$t_n = A_n T \qquad (2)$$

$$A_n = \frac{Total\ time\ taken\ to\ model\ all\ the\ same\ features}{Total\ time\ taken\ to\ complete\ the\ model} \qquad (3)$$

If the project contains "N" number of a specific feature (for example, multiple sketches), then the average time taken to complete a single feature (sketch) can be calculated as follows:

$$Average\ time = \frac{t_n}{N} \qquad (4)$$

Considering individual projects as small samples, the average time taken to complete a feature can be obtained from each one. For example, one might take "X" amount of time to complete a sketch for project A and "Y" amount of time for project B. All these sample averages can be clustered together and used in a statistical model such as Weibull distribution to estimate the population mean with certain confidence levels. This would give a best estimate as to how much

time on average one would take to complete a specific feature or task.

4.2.3 Statistical Modeling

The Probability Distribution Function (PDF) for the collected data (time) is unknown. As mentioned above, the chosen distribution should be dynamic enough to allow shape changes and predict parameter values. One such versatile distribution is the Weibull Distribution [66].

4.2.3.1 Weibull Distribution

Weibull distribution, first introduced by physicist Waloddi Weibull, has the capability to model various types of distributions simply by changing parameters. In the three-parameter Weibull:

Probability Density Function is given by:

$$f(t) = \frac{\beta}{\theta}\left(\frac{t-\gamma}{\theta}\right)^{\beta-1} e^{-\left(\frac{t-\gamma}{\theta}\right)^{\beta}} \quad \textbf{(5)}$$

Cumulative Distribution Function is given as:

$$F(t) = 1 - e^{-\left(\frac{t-\gamma}{\theta}\right)^{\beta}} \quad \textbf{(6)}$$

Mean is calculated using:

$$\mu = t_0 + \theta\Gamma\left(1 + \frac{2}{\beta}\right) \quad \textbf{(7)}$$

and Variance can be found by:

$$\sigma^2 = \theta^2 \left\{ \Gamma\left(1 + \frac{2}{\beta}\right) - \left[\Gamma\left(1 + \frac{2}{\beta}\right)\right]^2 \right\} \quad \textbf{(8)}$$

Where:

t: Random variable

θ: Scale parameter

β: Shape parameter

γ: Location parameter

Γ: Gamma function

t_0: Shifts the mean on the t axis

and:

$$f(t) \geq 0, t \geq \gamma$$

4.2.3.2 Parameter Calculations

There are numerous methods available to calculate the Weibull parameters; among them, the Maximum Likelihood Estimation (MLE) method is known to have yielded the best results, and was chosen for this research through a Python module. This method is derived from solving the following three non-linear equations [67].

$$\frac{n}{\beta} + \sum_{i=1}^{n} \log\left(\frac{t_i - \gamma}{\theta}\right) - \sum_{i=1}^{n} \left(\frac{t_i - \gamma}{\theta}\right)^{\beta} \log\left(\frac{t_i - \gamma}{\theta}\right) = 0 \quad \textbf{(9)}$$

$$\textbf{(10)}$$

$$-\frac{n\beta}{\theta}+\frac{\beta}{\theta}\sum_{i=1}^{n}\left(\frac{t_i-\gamma}{\theta}\right)^{\beta}=0$$

$$-(\beta-1)\sum_{i=1}^{n}\frac{1}{t_i-\gamma}+\frac{\beta}{\theta}\sum_{i=1}^{n}\left(\frac{t_i-\gamma}{\theta}\right)^{\beta-1}=0 \qquad \textit{(11)}$$

$$\beta>0$$
$$\theta>0$$
$$-\infty<\gamma<+\infty$$

4.2.4 Why Weibull Distribution?

Weibull distribution is capable of modelling symmetric data such as normal distribution and skewed data like exponential distribution simply by changing the shape parameters. The following figure shows how the changes in shape parameters change the distribution where the location parameter γ is assumed to be zero. Followed by that, Figure 16 shows that for specific parameter values Weibull distribution can be symmetrical.

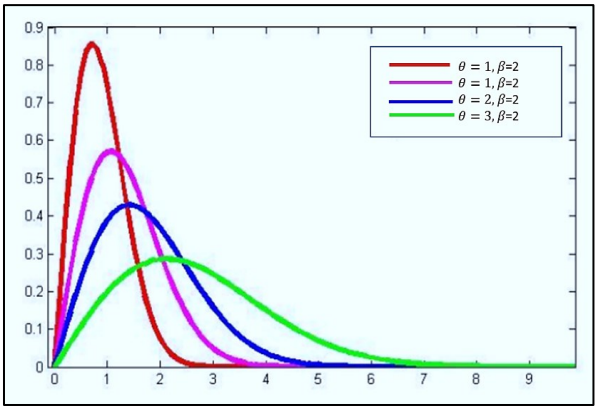

Figure 15: Weibull Density Plots for Varying β and θ

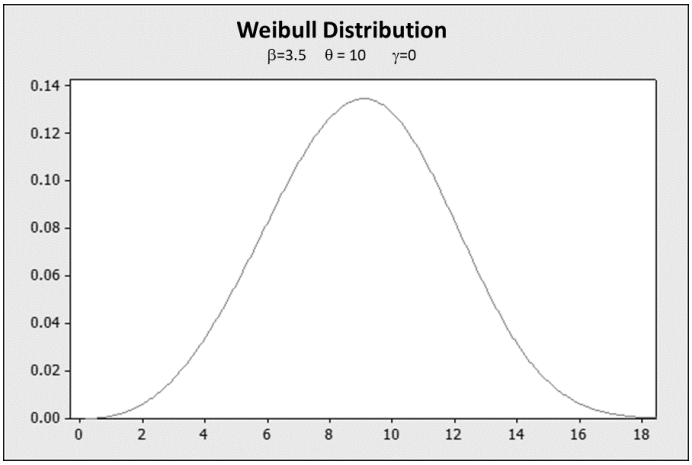

Figure 16: Symmetric Weibull Distribution

To this day, popular Weibull distribution applications are found in reliability engineering, failure analysis, wind-speed data analysis, unemployment duration, survival data [68], and PERT (Program Evaluation and Review Technique) [69]. The use of Weibull distribution in PERT is similar to the one discussed here. However, the application of improving employee productivity using big data is where this research differs from every other available application.

4.2.5 Weibull Application

Predicting the time needed to complete a task is something that many companies are looking to optimize in their processes. PERT was developed in 1957 as a way for managers to evaluate the time and resources necessary to manage a project [70]. It allows them to determine the possible paths for a project to be completed; through evaluation of the tasks, the most efficient process can be determined. Since the inception of this research, the following three questions have typically been used to determine the task completion time:

- What is the fastest, best case for completion? (Optimistic)
- What is the slowest, worst case for completion? (Pessimistic)
- What is the average time for completion? (Average)

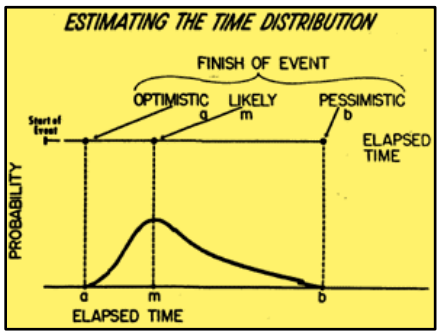

Figure 17: Estimating the Time Distribution

Although these questions have been asked over the years, finding the answers has been a difficult task due to the enormous amount of data and planning required to obtain the essential information. More recently, these questions were further researched by McCombs et al., where task analysis was performed using the Weibull Probability Distribution [70]. This distribution was chosen instead of Beta Distribution due to better compatibility with data analysis programs commonly used in engineering, such as Excel. The other benefit of the Weibull Probability Distribution is that it can have a variety of shapes. There are two characteristics that the Weibull Probability Distribution can describe that make this a preferable method [70]:

1. As processes become more efficient through planning and practice, the average time and the optimistic time begin to converge.

2. Processes tend to be very sensitive to anomalies such as equipment breakdowns, environment (snow, rain), labour and resource shortages, etc. which can be characterized as the Weibull shape with the longer right tail probability.

In this paper, the parameters used to define the Weibull Probability Distribution can be defined as the optimistic, pessimistic, and average time for a task completion. The optimistic and pessimistic parameters are defined in terms of the upper and lower percentiles of the probability distribution. With the Weibull probability curve, the upper and lower percentile can be any number but fitting with optimistic and pessimistic task completion, for which the 1st and 75th percentile values were chosen. A typical PERT chart is shown below:

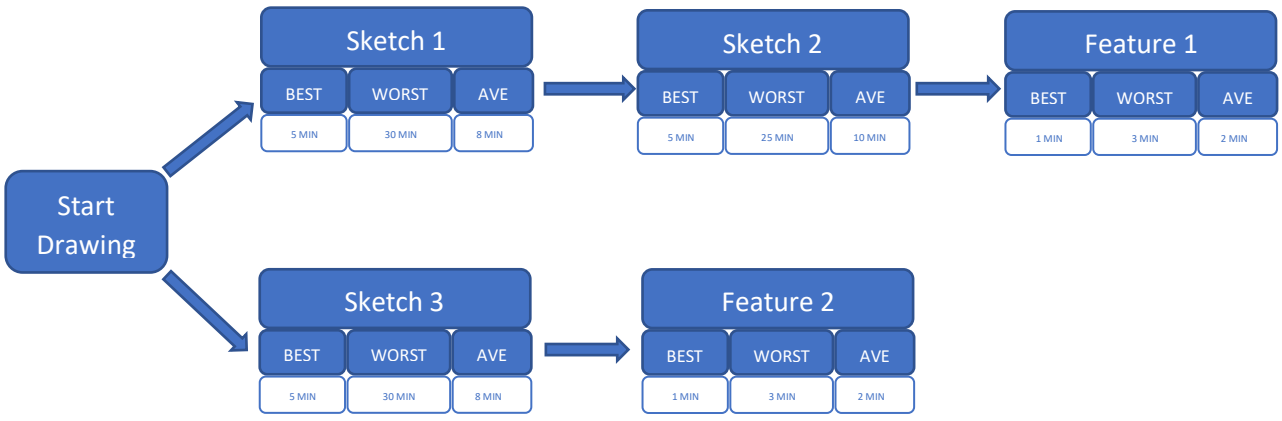

Figure 18: PERT Chart

It was also identified with the development of this process that understanding the details of the individual task was key to the success of this process. The PERT analysis, like the current research, would benefit from accurately understanding task completion times. The individual

tasks in the PERT chart were identified as having three typical completion times, as mentioned above: optimistic, average and pessimistic.

McCombs et al. chose using the Weibull Probability Distribution in situations like these since it is "…a distribution that can accommodate a longer tail probability than is allowed by the Beta Distribution" [70]. They took an existing study, "Defining a Beta Distribution Function for Construction Simulation" [71], and applied Weibull distributions to improve the outcomes.

4.2.5.1 Study on Construction Simulation

This study uses data collected from Vital Information Management System and Truck Payload Monitoring System that automatically records the performance data of a truck, such as time to load, distance, travel time, and time taken to unload. For this project, the data from 54,000 truck cycles were gathered using Truck Payload Monitoring System. The authors use Beta Distribution in this study because of its flexibility; just like Weibull distribution, Beta Distribution's shape can also be changed by changing parameter values. However, the best estimate for these parameters depends on the availability of performance data. In cases where these data were not available, the researchers in this study used subjective information gained from the skilled contractors [71]. To define a Beta curve unique to a specific case, one needs to define the minimum and maximum task completion times along with other statistical parameters such as mean, mode, certain percentiles and variance. In this case, the percentile characterizes "the percent chance that the analyzed activity duration will not exceed a time 't'" [71].

Authors in this study used VIBES (Visual Interactive Beta Estimation System) to estimate the Beta Distribution parameters. VIBES utilizes a combination of the following four characteristics: "mean and standard deviation," "mean and a selected percentile," "mode and selected percentile," and "two selected percentiles" [71]. The combination used in this study for an earthmoving operation is mode and 75th percentile with two endpoints. With this, the travel time approximation for an unfilled CAT 785 that moves at a range of 2.4-2.5 miles can be calculated with the help of VIBES [71]. This process requires the following subjective input information: minimum (7.5 minutes), maximum (two times the mode, 18.7 minutes) and most likely duration (9.3 minutes) [71]. The authors also assumed that there is a 75% chance the task duration would not go above 10.6 minutes (that is, 1.13 times more than the mode) [71]. With this information as input, VIBES produced the following graph and parameter values [71].

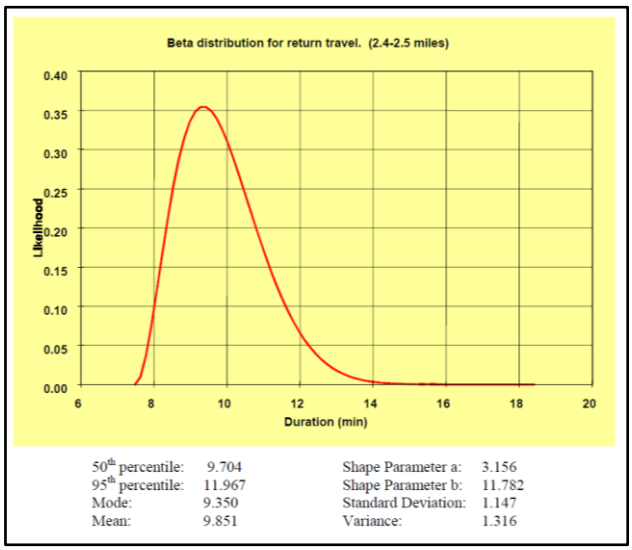

Figure 19: Beta Distribution for Return Travel

As can be seen, the calculated mean and variance in this study are an estimate of the real values. In addition, "The maximum travel time is computed as two times the mode. This assumption is supported by the reasoning that management would notice the slow-moving truck and take actions necessary to reduce its travel time. Undoubtedly, this type of assumption is necessary when a decision maker is constrained to using the Beta probability distribution" [70].

4.2.5.2 Task Duration Estimate Using Weibull Distribution

McCombs et al. explores the use of Weibull distribution in the case of above study [70]. However, instead of just using the traditional Weibull distribution, they introduce a new methodology that could be offset to accommodate tasks that have an optimistic task completion time greater than zero. In the paper, the three Weibull parameters (equation 5) are translated into three typical parameters that can be described as the optimistic (lower percentile), likely (mode) and pessimistic (upper percentile) values of the task completion time. The chosen lower percentile used was the 1st percentile and the upper percentile was the 75th percentile, although any percentile could be used [70]. The calculation of lower expert judgment percentile estimate (x_a) and upper expert judgment percentile estimate (x_b) are calculated using the following equations:

$$x_a = \theta[\ln(\frac{1}{R(x_a)})]^{1/\beta}$$
$$x_b = \theta[\ln(\frac{1}{R(x_b)})]^{1/\beta} \quad \textit{(12)}$$

Traditionally, Weibull distribution only requires estimates for two parameters out of the three. If the above x_a and x_b are known, then the Weibull parameters beta and theta can be calculated using the equations given by the authors, as shown in **Appendix B**. In addition, the threshold value for distribution can be offset to account for zero probability before a specific time using the following equations:

$$f(x-x_0)=\frac{\beta}{\theta}(\frac{(x-x_0)}{\theta})^{\beta-1}\exp[-(\frac{(x-x_0)}{\theta})^{\beta}]$$

$$R(x-x_0)=\exp[-(\frac{(x-x_0)}{\theta})^{\beta}] \qquad \textit{(13)}$$

$$M=x_0+\theta(1-\frac{1}{\beta})^{1/\beta}\text{ for }\beta>1$$

$$x=\theta[\ln(\frac{1}{R(x_b)})]^{1/\beta}+x_0$$

This threshold value would not affect the shape of the distribution; rather, it would just shift its position on the x-axis. Since the distribution is shifted the left boundary, information can be incorporated with the help of the equations given by the authors found in **Appendix B** [70].

4.2.5.3 Using Weibull Distribution in Construction Simulation

In order to use this method, only two percentile approximations and the mode approximation are needed, since the 75th percentile is already stated in the study [69]. The lower boundary is set as 7.67 minutes (representing the 0.01 percentile). The parameter values calculated are given

in **Appendix B** [70].

The figure below shows the Weibull distribution plot along with the initial Beta distribution.

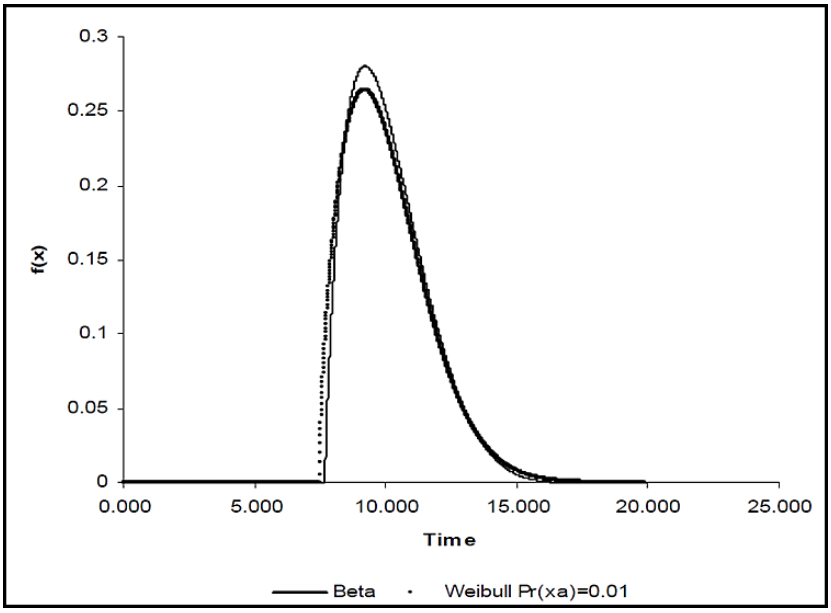

Figure 20: Plot of the Truck Travel Time from Fente, et al. (2000) and the Weibull Model [70]

The Weibull model is beneficial in this case because:

- It needs only three estimates as opposed to four.
- It is easy to model.
- It does not require an estimate of variance. It can be calculated precisely.
- The only potential error with the Weibull model stems from the initial human estimates.

The Weibull model, as mentioned above, is a versatile distribution that can be used to more accurately model task time distribution as long as the task time estimates themselves are accurate. This research explores a methodology which will eliminate any human estimation, allowing the

Weibull model to calculate parameters and distribution shape with greater accuracy, as discussed in later sections [70].

4.2.6 Case Study

A tracker program was developed by Enigma Design Solutions in Python and installed on company computers to track every activity of the user. Time spent on a specific project is recorded every second and data is saved into a MySQL/SQLAlchemy database. Collected data has the following attributes:

- Type of program;
- KEevents: tracks when there is a keyboard event such as typing;
- MEvents: tracks the mouse usage (such as when reading and scrolling through websites);
- SEvents: tracks every time the user scrolls.

The above attributes are utilized in data collection to avoid idle time. For example, if there is a document open on the computer and nothing is being done on it, then the program would not record the time.

Once the events are recorded individually, the total time spent on any project can be calculated in seconds by adding the event count. This can then be utilized to calculate the time taken to

complete each component of the project using time a weighted average method, which can then be fed into a statistical model. The figure below shows the workflow of this process.

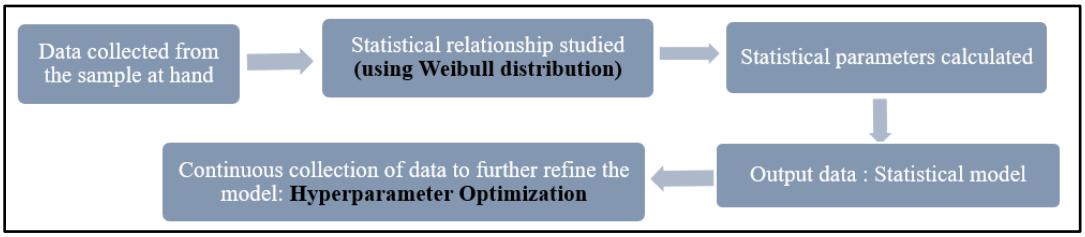

Figure 21: Workflow of Model Development

Data for the hyperparameter (time) by which the learning process is controlled is continuously collected, and the model is dynamically updated to refine the parameter values.

The figure below shows a typical dataset that is collected using the tracker.

sult Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content: | Fetch rows:

id	GUID	Program	WindowTitle	Win	Win	TimeStart	KEvents	MEvents	SEvents	TEvents	Activity	FDuration	Tags
3009123		python.exe	AE-0218043 - Walk Way Access Modification			2018-04-02 21:41:41	0	2	0	2	2	1	NULL
3376387		explorer.exe	AE-0218043 (International Paper (Grande Prairie) - Walk Way Access Modification)			2018-05-22 13:08:08	0	5	0	5	5	1	NULL
3376391		explorer.exe	AE-0218043 (International Paper (Grande Prairie) - Walk Way Access Modification)			2018-05-22 13:08:13	0	4	0	4	4	1	NULL
3609293		SLDWORKS.exe	SOLIDWORKS Premium 2018 x64 Edition - [AE-0218043 - Drawing - Sheet1]			2018-07-03 00:01:46	1	1	0	2	2	1	NULL
3609294		SLDWORKS.exe	SOLIDWORKS Premium 2018 x64 Edition - [AE-0218043 - Drawing - Sheet1]			2018-07-03 00:01:48	0	2	0	2	2	1	NULL
3609295		SLDWORKS.exe	SOLIDWORKS Premium 2018 x64 Edition - [AE-0218043 - Drawing - Sheet1]			2018-07-03 00:01:49	0	3	0	3	3	1	NULL
3609299		SLDWORKS.exe	SOLIDWORKS Premium 2018 x64 Edition - [AE-0218043 - Drawing - Sheet1 *]			2018-07-03 00:01:54	0	1	0	1	1	1	NULL
3609300		SLDWORKS.exe	SOLIDWORKS Premium 2018 x64 Edition - [AE-0218043 - Drawing - Sheet1 *]			2018-07-03 00:01:55	0	1	0	1	1	1	NULL
3609301		SLDWORKS.exe	SOLIDWORKS Premium 2018 x64 Edition - [AE-0218043 - Drawing - Sheet1 *]			2018-07-03 00:01:58	0	1	0	1	1	3	NULL
3609302		SLDWORKS.exe	SOLIDWORKS Premium 2018 x64 Edition - [AE-0218043 - Drawing - Sheet1 *]			2018-07-03 00:01:59	0	1	0	1	1	1	NULL
3609303		SLDWORKS.exe	SOLIDWORKS Premium 2018 x64 Edition - [AE-0218043 - Drawing - Sheet1 *]			2018-07-03 00:02:00	0	2	0	2	2	1	NULL
3609366		SLDWORKS.exe	SOLIDWORKS Premium 2018 x64 Edition - [AE-0218043 - Drawing - Sheet1 *]			2018-07-03 00:03:15	0	2	0	2	2	1	NULL
3609367		SLDWORKS.exe	SOLIDWORKS Premium 2018 x64 Edition - [AE-0218043 - Drawing - Sheet1 *]			2018-07-03 00:03:16	1	1	0	2	2	1	NULL
3609368		SLDWORKS.exe	SOLIDWORKS Premium 2018 x64 Edition - [AE-0218043 - Drawing - Sheet1 *]			2018-07-03 00:03:17	1	0	0	1	1	1	NULL
3609369		SLDWORKS.exe	SOLIDWORKS Premium 2018 x64 Edition - [AE-0218043 - Drawing - Sheet1 *]			2018-07-03 00:03:22	6	3	0	9	9	5	NULL

Table 3: Collected Data

For this data to be associated with a specific project, it is important for the window to be titled with the project name. The data gathered in this specific case is for an engineering firm. The data analysis model developed would focus on optimizing engineering projects and predicting engineering project times. This, however, can be generalized and modified to other fields of

work.

4.2.7 Data Sorting and Statistical Modelling

As mentioned above, the data collected is based on time spent on a specific project. For example, the figure above shows the time spent on a project that involves Solidworks modeling and analysis. This then can be sorted further based on the different components of that project to study the time spent by the user on those components.

Making an engineering model in Solidworks takes several steps. A VBA macro is developed and implemented to extract the details on those steps involved in completing the Solidworks model and store them in an Excel file.

The following table shows an example of the data extracted from a Solidworks model shown in the figure using the VBA macro. This can be done for every part that is modeled in Solidworks for a specific project.

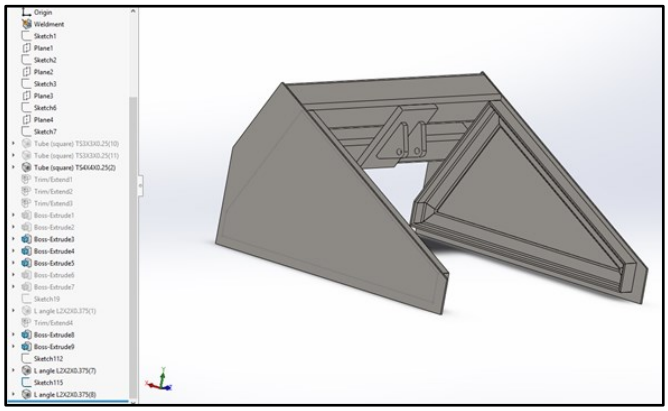

Figure 22: Solidworks Model

Comments	Front Plane	Design Binder
Favorites	Top Plane	Annotations
History	Right Plane	Surface Bodies
Selection Sets	Origin	Solid Bodies
Sensors	Imported3	Lights, Cameras and Scene
Sketch2	Sketch1	Sketch3
Cut-Extrude2	Cut-Extrude1	Cut-Extrude3
Equations	Sketch4	Material <not specified>

Table 4: Extracted Data

The Excel data collected is then transferred into a MySQL database through a program developed in SQLAlchemy for each model. The breakdown of this information is then used as data for the time variable and trains the statistical model.

The table below shows how the Solidworks data is transferred into the MySQL database for each project.

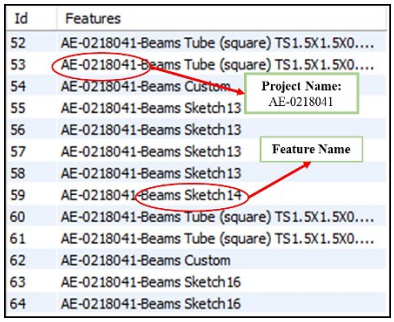

Table 5: Transferred Data

As shown above, **Project AE-0218041** is an engineering project that mainly involves solid modelling in Solidworks, as well as FEA analysis. Solidworks modelling includes feature modelling such as drawing sketches and extruding. From the data extracted, the total number of

instances for each feature is calculated as shown in the table below.

id	Title	sketch	extrude	weldment	tube	Total_Time_Spent	Total_Number_of_Instances	Average_Time
45	AE 0218041 Spacer	32	20	2	66	675	140	4
46	AE 0218041 Base Plate	23	10	0	0	36	45	0
47	AE 0218041 Beams	36	4	2	12	564	66	8
48	AE 0218041 Bench	196	12	8	56	2969	312	9
49	AE 0218041 Bench2	36	0	2	16	0	56	0
50	AE 0218041 Bench3	36	0	2	16	32	56	0
51	AE 0218041 Bench4	8	0	0	0	4	10	0
52	AE 0218041 Bench5	78	12	2	8	2144	132	16
53	AE 0218041 Channel	26	4	2	0	156	36	4
54	AE 0218041 Columns	14	4	2	4	261	30	8
55	AE 0218041 EndCap	12	4	0	0	421	20	21
56	AE 0218041 Gusset	4	4	0	0	225	12	18
57	AE 0218041 Lug	130	47	2	2	10598	244	43
58	AE 0218041 Mount Plate	52	29	0	0	1533	110	13

Table 6: Features Summary

Then, the total time spent on each project is calculated using the data obtained from the tracker. Using the total time and the total number of features, the Time Weighted Average is calculated for each feature as shown in the figure below. This then is used to calculate the average time taken to finish one of such features.

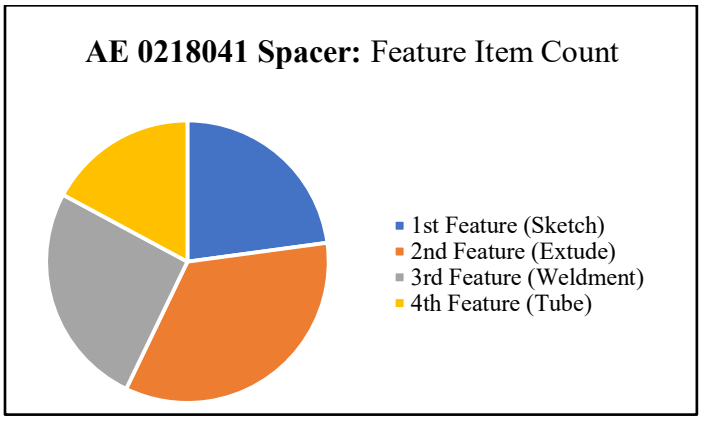

Figure 23: Item Count

This can be done for any project, allowing all the data to be clustered together to form a sample of average times for each feature. This is then fed into a statistical model based on Weibull distribution to determine the shape and parameters in Python. The figures below show the

calculated Weibull parameters for two different features. It also shows how those parameters are changing as more data is added to the distribution. The addition of data changes the parameter values of Beta and Eta, as well as the shape of the distribution changes. As shown below, a higher Beta value gives normal distribution and a value less than one produces an exponential distribution.

Feature Name: Sketch

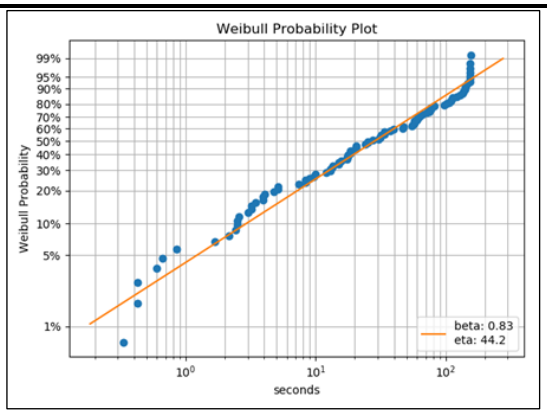

Figure 24: Weibull Probability Plot (I) for Feature Sketch

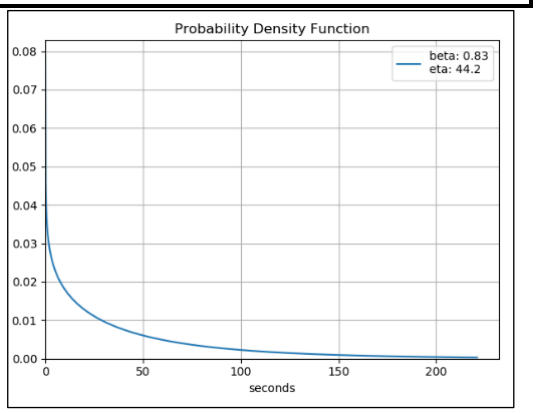

Figure 25: Probability Density Function (I) for Feature Sketch

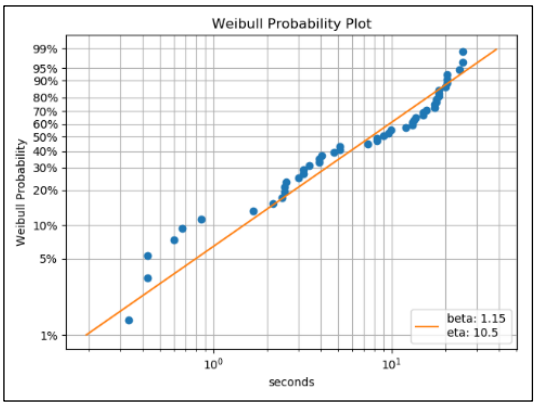

Figure 26: Weibull Probability Plot (II) for Feature Sketch

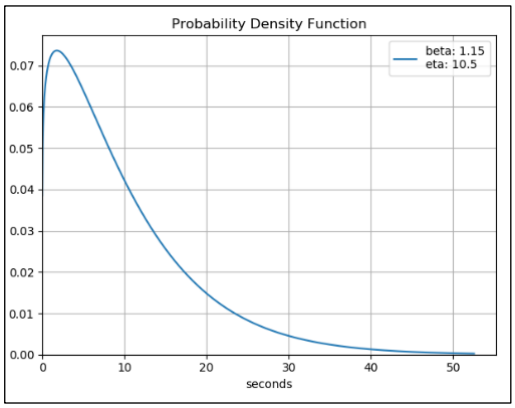

Figure 27: Probability Density Function (II) for Feature Sketch

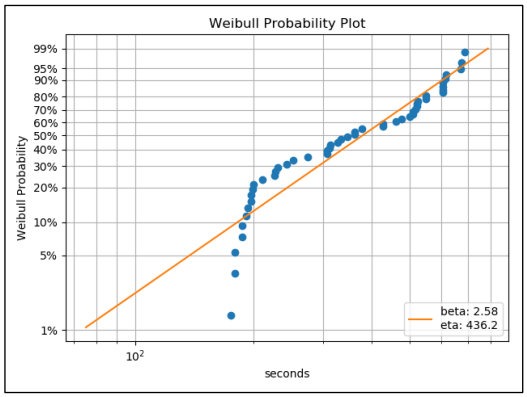

Figure 28: Weibull Probability Plot (III) for Feature Sketch

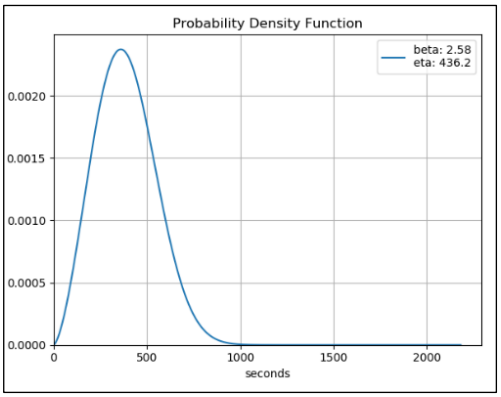

Figure 29: Probability Density Function (III) for Feature Sketch

Feature Name: Extrude

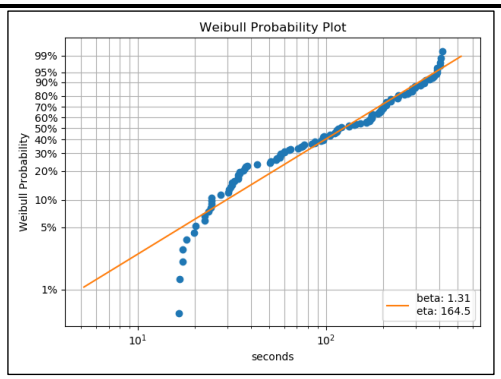

Figure 30: Weibull Probability Plot (I) for Feature Extrude

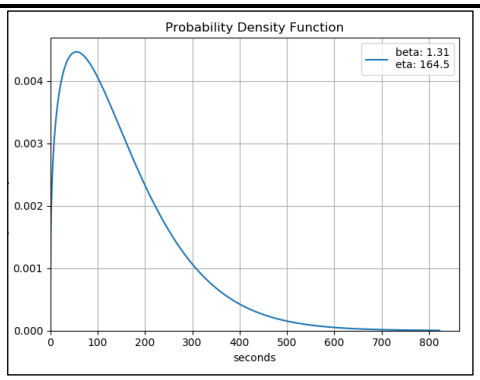

Figure 31: Probability Density Function (I) for Feature Extrude

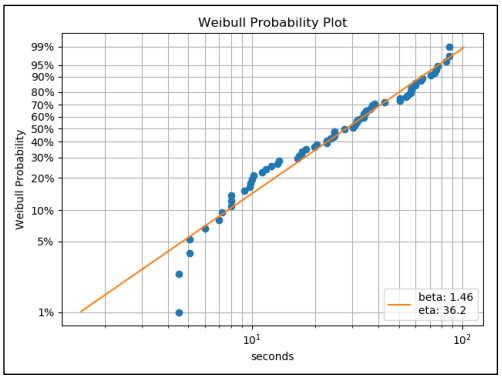

Figure 32:Weibull Probability Plot (II) for Feature Extrude

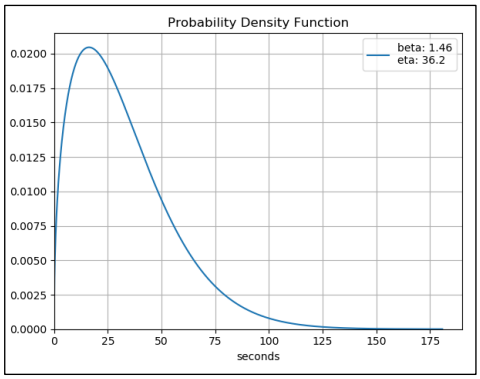

Figure 33:Probability Density Function (II) for Feature Extrude

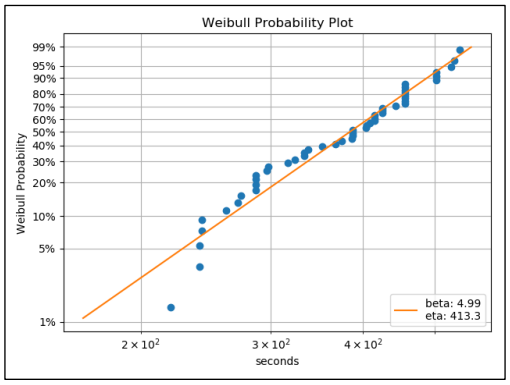

Figure 34: Weibull Probability Plot (III) for Feature Extrude

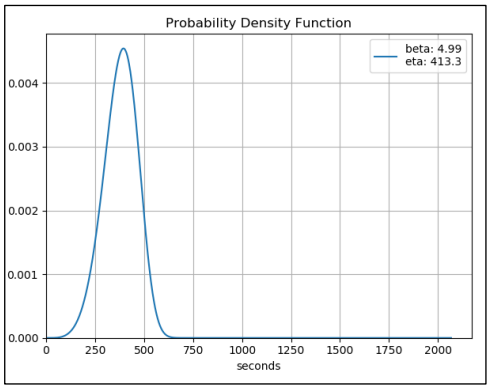

Figure 35: Probability Density Function (III) for Feature Extrude

Dependency and Concurrency: When considering this model, one must also consider the interrelationships between various tasks or features. Most of the tasks performed by employees depend on previous or concurrent tasks. Therefore, the time taken for a specific task would depend on how well the tasks before and after it are done. An example of such a case is shown below; as can be seen, the time taken for an intermediate step depends on how well the previous one was done.

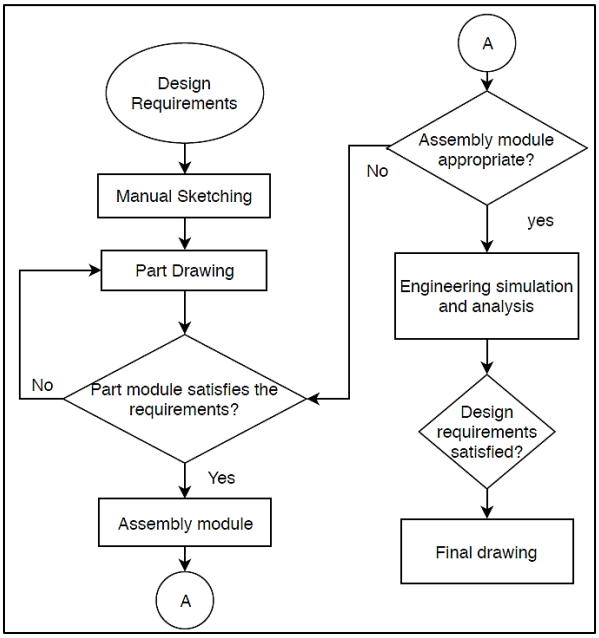

Figure 36: Dependency Diagram

The chosen statistical model provides a range of values a specific task would take instead of just giving an exact value. As mentioned earlier, the sample averages are obtained to model the probability distribution, which then is used to estimate the population mean with a preferred confidence level as shown below for normally distributed data.

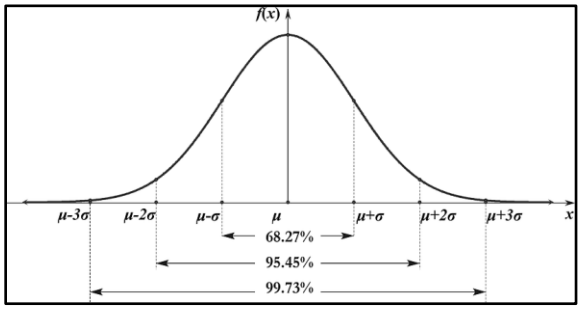

Figure 37: Confidence Level: Normal Distribution

Depending on how spread-out the data are, one could analyze the factors affecting the shape of the distribution. For example, if the distribution is too spread-out, it may be due to

- Extreme interdependency
- Employee is in training
- Interdependent tasks are being delayed

On the other hand, if the distribution is narrowly shaped, then the opposite of what is mentioned above can be assumed. An example of such varying distributions is shown below.

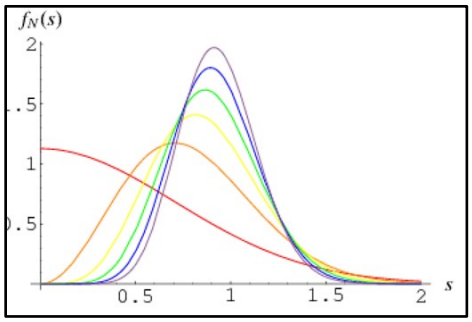

Figure 38: Various Distributions

Based on this, one could analyze how a specific task is being influenced and how it could be improved. For the future, the tracker program could be improved to track the time spent on not just the overall task, but also every intermediate task.

4.3 Discussion

With the data gathered and the statistical model in place, how can these methods benefit

companies and employees? As explained in the introductory paragraphs, each actor in today's economic environment faces different types of challenges. How can the above model be the first step in using big data to mitigate all those issues? Let's look at how each actor individually can benefit from this solution.

4.3.1 Businesses

Optimization and human capital management: Human capital optimization is essential for businesses to increase their productivity, profit, and customer satisfaction (employees are also considered a customer here). Knowing employees' strengths and weaknesses would help a company utilize the right talent in the right place, just like how engineers choose design solutions based on the problems at hand. The tools they choose to design the solution are based on the technical requirements; for example, material selection is based on the material properties and the design requirements such as strength, type of loading, or temperature. The availability of material data makes it easier to choose the appropriate material in just few seconds. Imagine being able to do the same with human capital: knowing who is good at what and how much time each individual takes to complete tasks when putting together a team for a project. That can be done using the above model.

The above statistical model would also allow for an accurate current task time estimate with a specific confidence level. It is current because the model is dynamic. This estimation can then be utilized to

- Accurately estimate project completion time;
- Predict exact cost of human capital;
- Optimize human resource utilization.

Performance evaluation/management: "Performance management (PM) is a goal-oriented process directed toward ensuring that organizational processes are in place to maximize the productivity of employees, teams, and ultimately, the organization" [72]. Traditionally, performance management is done through performance appraisal, "…a formal system of review and evaluation of individual or team task performance," done by managers on an ongoing basis. The following are some of the modern evaluation methods that are being widely used in the industry.

- **Multisource (360-Degree) Appraisal Technique:** Done by multiple stakeholders such as peers, customers, supervisors, and subordinates [73].
- **Management by Objectives (MBO):** Evaluated against goals set by management and achievement by the employee [73].
- **Behaviorally Anchored Rating Scales (BARS):** "BARS contrast an individual's performance against specific examples of behavior that are anchored to numerical ratings" [73].
- **Assessment Center:** Where behaviours are observed in set exercises as an evaluation method [73].

The above methods of performance evaluation are subjective and often not very positive. A majority of employees dislike performance appraisal; a recent survey shows that 80 percent of employees expressed displeasure with their performance evaluation system [74]. Performance evaluations are usually done for the following reasons [74]:

- Recruitment
- Training and career planning
- To set compensation programs
- To assess the potential of an employee

As shown, important decisions about a company's human capital management are based on subjective evaluation. With the above-mentioned model, companies can instead use hard data to build a system to evaluate an employee's ability to perform tasks, as well as the actual contributions he/she makes to the company.

Training and local skill development: Instead of moving to a different part of the world to find skilled workers, companies can use the above solution to train new employees and help improve the skills of the existing ones.

Offer flexibility to employees: Being able to track and record work would help companies offer a flexible work environment, allowing employees to work from anywhere and at any time. This is something the new generation of workers expects and is attracted to; having this flexibility would help companies draw in and retain talent.

4.3.2 Employees

Self-evaluation: This would help identify weaknesses and strengths. Employees would be able to understand the contributions they are making to the project. This self-evaluation would be purely based on data collected rather than a subjective performance evaluation.

Virtual mentor: A mentor is someone who shares knowledge, gives advice based on his/her career experience, and offers guidance when needed to help mentees succeed in their chosen profession. As millennials are taking over the workforce and learning the ways of industry, it is very important for them to have the right mentors to guide them. Whether the industry or the training institutions want to admit it or not, there is a gap between what the industry requires and what training institutions can offer. This gap would have to be filled by the senior managers/mentors in the industry. However, as mentioned above, industries do not always have that capability under fragile economic conditions. The above model can act as a virtual mentor and guide an employee to understand their weaknesses, strengths, and contributions; this will help them improve themselves or use their strengths to build on their careers.

4.4 Disadvantages

One might ask why an employee would willingly sign up to have Big Brother watch over their shoulder and allow the organization to scrutinize every second they spend while being paid. In order to solve that problem and encourage employees to take advantage of this solution,

companies can:

- Give ownership of data to the employee and allow them to share only what they want
- Reward those who share their data and find ways to improve themselves through gamification of work
- Emphasize that it is there to help them rather than monitor them

4.5 Future Work & Generalized Model

The above model can be generalized for different types of tasks and various firms, not just engineering firms. Using the new technology available, the tracking can be done not just for those who work on computers, but also other types of workers. There are tracking devices that could collect data using the tracker (used in this research), GPS, internet, and every other digital medium available. These data can then be fed into the statistical model. In addition, a mathematical optimization model could be developed to help with project management using this dynamic data model.

5 Gamification Model

The application of technological game elements in non-gaming contexts to harness certain human behavioural elements that produce preferred results is not a new phenomenon. It is the wide-ranging popularity and the acceptance of computer games that led to the increased use of game elements in non-gaming industries, and it continues to gain momentum as a business

concept in the 21st century. The gaming industry's worth has seen continuous growth in the past six years and is projected to reach $90.1 billion by 2020 from $54.20 billion in 2011 [3]. This growth has been facilitated by extensive research on human motivations in game playing, as well as technological advances that enabled researchers to directly apply new theories to practice. This report discusses the theories and the current applications of gamification and explains a modified gamification model developed by us, which uses real time data and a machine learning algorithm, along with a new rewards system.

5.1 Case Study: Effects of Simultaneous and Sequential Work Structures on Distributed Collaborative Interdependent Tasks

Statistical analysis has been used in the past to understand various work structures. This section looks at one such example that uses statistical modelling to understand the effectiveness of sequential and collaborative work structures on interdependent tasks as a function of number of workers. The hypothesis tested in this study is: groups perform better in interdependent tasks through a sequential work structure rather than a simultaneous one, while reducing the process losses [75].

The experiment takes place in the context of Amazon's Mechanical Turk, a crowdsourcing website that helps companies find workers to do on-demand tasks. Users are asked to perform a creative task: writing a limerick [75].

Experiment setting and analysis summary [75]**:**

- A two by two experiment, where participants followed a sequential and simultaneous task flow.
- Groups contained one, two and three participants.
- Baseline condition is one-person group to which the two- and three-member groups are compared.
- Simultaneous task: groups of two or three workers worked simultaneously online.
- Sequential task: workers did tasks in iteration adding additional workers one after the other.
- Results were tested on two effects:
- Task condition (simultaneous or sequential).
- Interaction effect through work quality as a **2 by 2 ANOVA**.
- Outcome is rated based on technical quality such as number of lines, rhyme structure, meter, and quality dimension such as creativity, coherence, etc.
- Process measures considered: time spent, characters written and deleted, edits' distribution, and number of author changes.
- Three people rated the limericks for quality and technical dimensions based on 1-7 scale.

- **Cohen's alpha** is calculated and used to compare the dimensions considered to test their independence. The ones with lower alpha values were combined to form one scale.

Results and Analysis [75]**:**

The table below shows the overall result of the experiment.

Number of Worker	Simultaneous			Sequential		
	Mean	SE (Standard Error)	n (number of groups)	Mean	SE	n
1	4.05	0.1	79	4.05	0.10	79
2	4.03	0.09	107	4.48	0.07	78
3	4.15	0.11	89	4.59	0.08	78
combined	4.08	0.06	275	4.37	0.05	235

Table 7: Overall score of final limerick. Simultaneous workers (group sizes of 1, 2, 3 workers), and Sequential work (iterations of 1, 1+1, 1+1+1 workers) [75]

The authors performed a 2X2 statistical analysis, removing the one-worker groups as a baseline and concluding that the "…mean quality for simultaneous and sequential conditions differ, and the effect is similar in groups of two or three workers. Larger groups (or groups with more steps) don't differ from smaller groups" [75].

Following that, a comparison of one-worker groups to multi worker groups concludes that "Simultaneous groups of 2 and 3 workers (M=4.08, SD=0.99) did not produce higher quality limericks than individual workers." On the other hand, "… sequential groups of 2 and 3 workers (M=4.53, SD=.11) produced higher quality limericks than individual workers (M=4.05, SD=.93)" [75].

This is an example where statistical analysis is used to evaluate quality of work in two

different settings. As mentioned above, the factors and evaluation method considered to study the quality were initially set by the researchers in order to assign a specific scale. Similar to that, real time data, statistical analysis, and machine learning are used in our research to develop a gamified model that completely eliminates human subjectivity (such as the three raters used in this case study) in evaluation that in turn is used in a chosen model to assign the points earned by the users.

5.2 Model Development

Changing technology and workplace demographics require companies to find innovative ways to keep their employees engaged and productive. Research shows that the expectations of the new generation that is currently entering the workplace are drastically different from those of their predecessors. Companies that are trying to adapt to the changing expectations of their workers tend to do better than the ones that are reluctant to change [76]. The model developed is an employee-focused dynamic game model that takes all the motivational factors above into consideration. The sections below discuss the model development in detail, as shown in the overall diagram below, and a reward system that would trigger both intrinsic and extrinsic motivation.

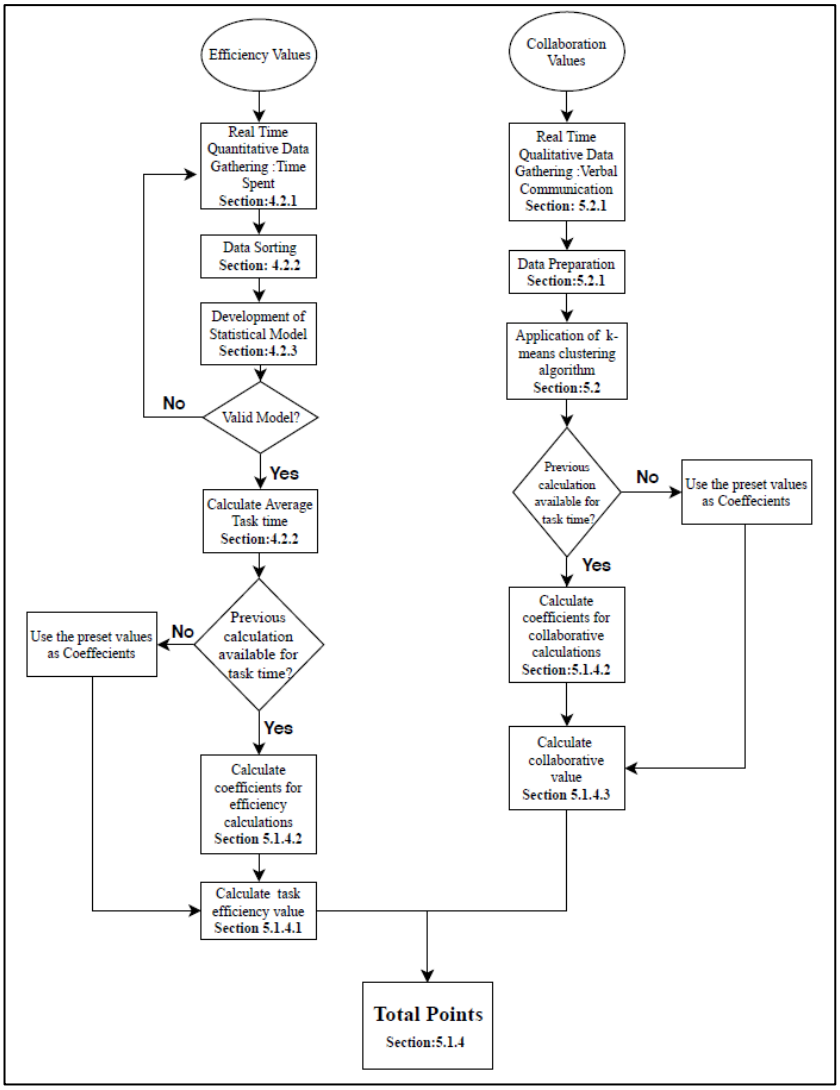

Figure 39: Overall Model Development Diagram

5.2.1 Data Collection

In the model, real-time data is collected through an integrated tracking software that tracks a user's every task. Examples of such activities for a small inspection/engineering company are

given below:

- Drafting
- Marketing analysis
- Testing
- Design analysis
- Inspection of a 50-ton crane
- Safety inspection
- Writing of an inspection report
- Engineering calculations

With the data gathered for each of the above-mentioned activities, a statistical model using Weibull Probability Distribution can be built to find the average time taken by a worker to do a certain task, as shown below.

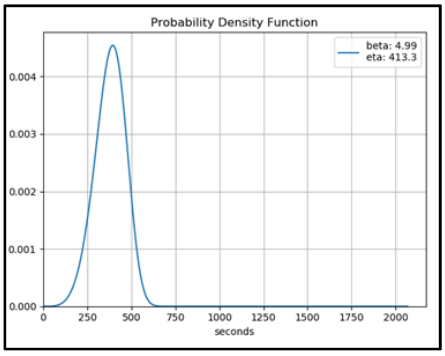

Figure 40: Example of Weibull Probability Distribution and Parameter Estimation

In addition to work-related tasks, the tracking device can also track every activity of the user by gathering data from other organizational and social applications such as GPS, Facebook, fitness trackers and email.

5.2.2 Factors Considered

The type of game design the current work environment needs should include elements that allow cooperation among employees for the following reasons:

- The generational changes in workplaces
- More diversity in the workforce, with people of different backgrounds and points of view
- Increased multilocation project teams and virtual projects
- To help employees set measurable goals to achieve targets in the workplace
- To improve work quality through real-time data-based feed back
- To help employees understand their position in the organization and what they can do to advance their careers
- To help improve work-life balance through rewarding employees for a healthy lifestyle

Some benefits of employee cooperation are as follows [77]:

- Increased productivity
- Improved employee involvement

- Job satisfaction that would lead to employee retention
- Positive organizational culture

How can cooperation be promoted?

- Reward cooperation
- Encourage and reward collective achievement

Competitiveness should be allowed without turning the work environment toxic. However, this is unavoidable if the gamified point system is used to decide who advances in the organization, as discussed below in the rewards section.

Competitiveness could be allowed:

- Against past individual performance, as in a single-player game, with employees constantly competing to be a better version of themselves
- Against other teams rather than individuals
- As soft competition (e.g., who does well based on how well they use their rewards to improve their skills)

5.2.3 Dynamic Game Model

A dynamic or sequential game model is recommended in this case. According to Hespanha, dynamic games have infinite stages; action spaces are not defined by finite sets [78].

This concept is described by the following equations [78].

Multiplayer game:

$$x_{k+1} = f_k(x_k, u_k, d_k), \forall k \, \epsilon \, \{1,2, \dots K-1\} \quad \textit{(14)}$$

Where:

f_k: "Dynamics" at stage k

x_{k+1}: Entry node at stage k+1

x_k: Entry node at stage k

u_k: Player one's action at stage k

d_k: Player two's action at stage k

Single player game:

$$x_{k+1} = f_k(x_k, u_k), \forall k \, \epsilon \, \{1,2, \dots K\} \quad \textit{(15)}$$

Following the above concept, the developed model has infinite stages; it can be continuously played as long as the employee continues at the organization. The action space is defined by an infinite set, as there is no limit to what the player can choose to do. It is up to them to strategize and play the game to maximize their rewards based on the results they obtain, as

explained in the sections below.

5.2.4 Mathematical Modelling

Research has found the following factors to be the most influential or important ones to tackle some of the main challenges faced by companies and employees:

- Collaboration
- Efficiency

Other factors can be added into the model as needed depending on the environment. Assuming the earning points or credits are a function of the above factors and the relationship is linear, the function can be written as follows:

$$points = f(x_1, x_2) = ax_1 + bx_2 \quad \textbf{(16)}$$

Where:

a & b : Constants

x_1: Efficiency

x_2: Collaboration

5.2.4.1 Calculating x_1 $(efficiency)$

Productivity calculation methods are traditionally used in the manufacturing industry and are "…defined as a ratio of a volume measure of output to a volume measure of input use" [23].

Efficiency calculation in engineering, similar to productivity calculation, is done as the ratio between useful power output to total input. Besides using factors such as gross output, number of finished projects, and items sold, there should be a mechanism to track and evaluate not just the final output, but also all the tasks involved in reaching the final output.

For this research and the gamification model, the employees' efficiency is considered a function of time: task completion time, to be specific. The values for task completion time are obtained from the statistical model, as explained in the data collection section. The time taken to complete all the subtasks of a project is tracked and the average is calculated using the Weibull statistical model with a certain confidence level.

Work-life balance can be calculated based on how much personal time an employee spends outside of work to improve themselves, such as doing leisure activities and working out. Work-life balance is considered one of the major factors that affects employee productivity; therefore, tasks done outside of work should also be included in calculating efficiency [79]. With this, the efficiency calculations can be formulated as follows using a linear relationship:

$$Efficiency = x_1 = f(y_n) = \sum_{i=1}^{n} a_n \, y_n \qquad \textbf{(17)}$$

Where:

y_n : Average time taken to complete a task

5.2.4.2 Calculating a_n & b_n (Constants)

Average time here is calculated from the statistical model. Constant term is calculated as the ratio between the average times taken previously and the current average time. For example, if the employee took "Y" amount of time to complete a task for Project A and took "Z" amount of time for the same task for a previous Project B, then a_n can be calculated as follows:

$$a_n = \frac{Z}{Y} \qquad \textit{(18)}$$

Based on the time taken, the ratio would be less or greater than one; that would increase or decrease the value of efficiency. The benchmark used here to measure efficiency would be the employee's best average, rather than a pre-set value by the organization. In other words, employees compete with themselves to get a better performance value or efficiency score. b_n can be calculated in a similar fashion.

5.2.4.3 Calculating x_2 (Collaboration)

Traditionally, collaboration is measured through feedback and surveys. This, however, is subjective and can be affected by outside factors such as human emotions and perceptions. Big data and machine learning algorithm K-means clustering (unsupervised learning) can be used to calculate the range of collaboration among employees using their verbal communication; the scores calculated can be used as the value for x_2 . Further explanation on how this works is given in the section below.

Additionally, the nonverbal collaboration can also be calculated through collective

achievement, such as how well a team has done on a specific project based on the completion time, against budget numbers and quality. The methodology selected in the previous section to calculate individual efficiency can be used to calculate group efficiency against other groups and used as a value for collaboration.

5.3 Applying K-means Clustering

K-means clustering is an unsupervised learning algorithm that identifies similar data points in a given data set and groups them into "k" number of clusters as outlined below for a set of data points $x_1, x_2, x_3 \ldots\ldots x_n \in \mathbb{R}^d$. A K-means clustering algorithm begins with an initial group of arbitrarily selected centroids that are used as starting points for each cluster. Following that, the points are assigned to the closest center based on the Euclidean distance. Once the points are assigned to the clusters, the location of the centroid is recalculated based on the average distance of the assigned points. This continues until the optimal centroid location is achieved or the predefined iterations are reached. The figure below shows how the clusters and centroids are formed. Followed by that the steps in implementing this algorithm are given.

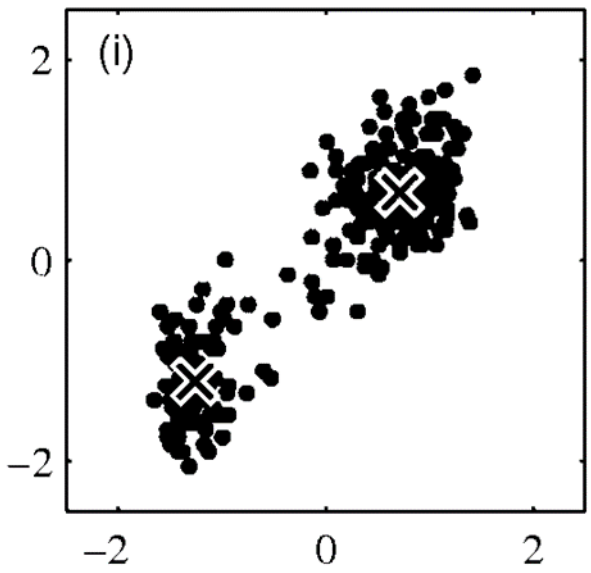

Figure 41: K-Means Clustering Illustration

- Select the "k" number of clusters to be identified for a set of raw data. The initial value for "k" could be anything, as it will be modified by the algorithm later to best fit the data.
- Place a centroid (mean position of all data points) denoted by $\mu_1, \mu_2, \mu_3 \ldots\ldots \mu_k \in \mathbb{R}^d$ at any location in the cluster to begin the process.
- Following that, the algorithm finds the nearest centroid for each point y_j and places it in that cluster. The algorithm aims to minimize the mean squared distance (μ_k) between the point and the centroid in this step as given by the following K-means cost function [80]:

$$f_{k-means} = \sum_{i \in [n]} \min_{j \in [k]} \left\| x_i - \mu_j \right\|^2 \qquad \textbf{(19)}$$

- The algorithm continues the iterations for different values of "k" (number of clusters) until the optimal value of "k" is found.
- An "elbow plot" that graphs the reduction in variation and the number of clusters (k) is used in deciding the optimal value for "k." For example, as shown in the figure below at k=3, there is a huge reduction in variation, and it does not go down as fast following that point.

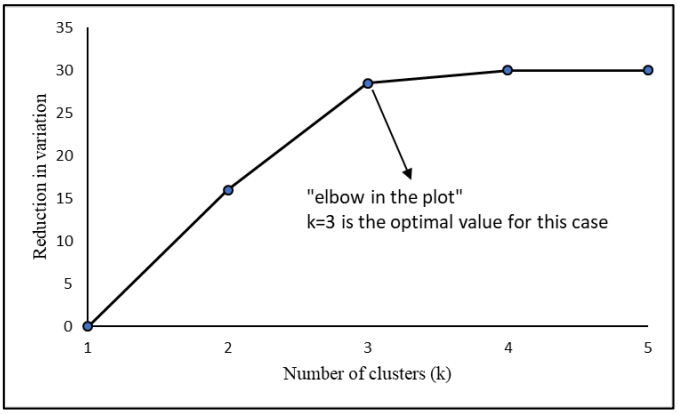

Figure 42: Reduction in variation vs. number of clusters (k)

5.3.1 Measuring Collaboration Using K-means Clustering

Using the K-means clustering algorithm, one of the ways collaboration could be measured is through tracking the written communication between employees and measuring the frequency

of certain keywords. Example of such keywords could be “commitment,” “flexibility,” “participation,” “trust,” and words that acknowledge others’ work, such as “thanks,” “great,” or “absolutely.” This act of processing and analyzing the human language is considered a type of Natural Language Processing (NLP). This group of tools and processes have been a part of ongoing research development since the 1950s [81]. With the advances in computer technology, the techniques and models have become more detailed and efficient. For this case, the data is conditioned for analysis using the following techniques:

- Standardize data, i.e., convert them all into lowercase
- Eliminate any punctuation, HTML, or stop terms
- Tokenize or divide text into meaningful groups
- Lemmatize terms that identify the lemma of a term
- Finally, transform the text into a vector of numbers using Term Frequency-Inverse Document Frequency (TF-IDF), which helps assess the relative importance of a specific term in a document.

Once the data is ready, following the above steps, the sklearn package in Python can be used to implement the K-means clustering algorithm to analyze the text data (written communication). It gives a general understanding as to the content of the text data that is being analyzed. For example, the most frequently used terms can be identified and given a score as

shown in the figure below.

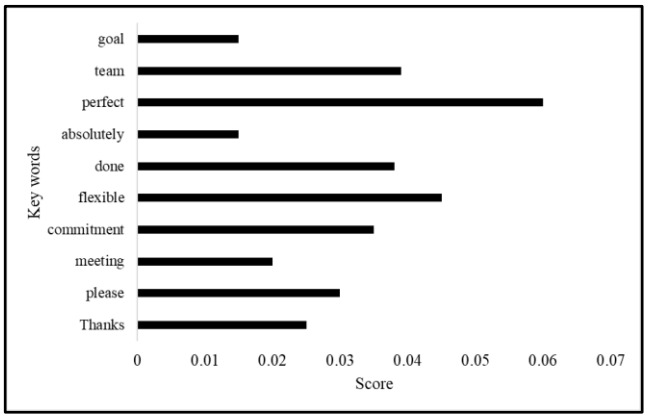

Figure 43: Keywords vs. Score

This alone would not be enough to evaluate the true intent of any verbal communication. These words should also be analyzed based on the context to have a clear understanding of the type of information that is being relayed. For a simple case, a word could have a positive or negative connotation. Other attributes, such as whether the word is passive, aggressive or a power word, can be other scales that should be modelled. An NLP processor could be utilized to rank the words in their context for use in efficiency calculations.

5.3.2 Using NLP Processors for work gamification

Soon after the unveiling of IBM's Watson on the game show *Jeopardy!*, an article was presented by Tony Pearson that explained the components of how Watson worked and provided examples of resources for each component, as shown in **Appendix A** [82].

This architecture would be similar for most machine learning applications, although the

tools required to achieve desired results would depend on how well the data can be mined and extracted.

Communication is one of the fundamental aspects of any relationship, whether it be with employees, coworkers, clients or the public. There are essentially three main types of communication:

- Visual Verbal (Meetings, Skype)
- Non-Visual Verbal (Phone, Voice Chat)
- Written (Letters, Text)

5.3.2.1 Visual Verbal Communication

This type of conversation is through spoken words, in either a face-to-face conversation or using other communication mediums such as smartphones or Skype. This type of communication allows people to convey much more than just spoken words. The tone, posture, eye contact, intonation, pitch variances and body language are some things that can be used to communicate a person's intent besides the words used. The same phrase can have completely different meanings based on how it is presented in visual verbal communication. This is considered the most effective type of human communication.

5.3.2.2 Non-Visual, Verbal Communication

This entails verbal communication with no visual components. An example of such communication could be non-visual telephone calls. It is a little more difficult to process this

type of communication, as no visual communication clues exist and the intent of the information that is exchanged is evaluated solely based on vocal tone, pitch and cadence.

5.3.2.3 Written Communication

This is the most difficult form of communication because the tone of the conversation is found only in the words and punctuation. As part of the recent industrial revolution, this form of communication has also evolved, thus creating two classes of written communication: formal written communication and informal written communication.

5.3.2.4 Formal Communication

Formal communication is defined as "…goal-oriented, explicitly stated, function related communication that flows through the hierarchy, follows prescribed norms, and transcends time and space"[83]. When analyzing these types of documents to provide insight into items likes content, relativity and tone, more sophisticated tools are required to ensure the accuracy of the results. In certain scenarios, such as company correspondences, the scope of document context becomes relatively repetitive. These communications are usually regarding items such as:

- Contract Negotiations
- Work Progress
- Performance Evaluations
- Memoranda
- Complaints

As documents posted to the company must evaluated on content, relativity and tone, it is more likely that a template will be used to set the standards for a specific type of document.

For the most part, the template would be created such that only a few basic details would have to be changed, such as names and dates. Depending on the situation, the success of the document would depend only on how well it conveyed the information. However, in some other cases, such as employee evaluations, the structure of the communication greatly affects the result and is thus different for each employee.

5.3.2.5 Informal Communication

This is a casual or unofficial method of communication, such as the communication between friends or peers. It can be either written or oral communication. In this paper, written informal communication will be considered for analysis. One of the first evolutions in written informal communications came as a result of the increase in text messaging on mobile phones before keyboards were integrated as part of the phones' architecture. Initially, the keyboard for text messaging was also the number pad. This resulted in three letters being associated with each number, where the number of times the button was pressed corresponded to the letter typed. For speed and simplicity, many common words and phrases were shortened to abbreviations. Some common abbreviations include the following:

Abbreviations	Full Form
lol	Laugh out loud
brb	Be right back

L8R	Later
wbu	What about you
u	You
r u	Are you

Table 8: Information Communication Abbreviations

As one can see, some of the abbreviations above are acronyms like those found in more traditional forms of writing, while others use shared phonetics to shorten words or sentences so that the messages make sense when read aloud. Some other nuances in informal communication include things like typing in all capital letters (all caps) to represent yelling. Having a period at the end of a sentence indicates the end of a discussion, while multiple periods at the end of a sentence indicates there is more to be said about the preceding sentence, but it will not be said directly.

In recent years, this form of communication has experienced an evolution in which keyboard symbols began to be combined in unique patterns to represent faces, emotions, or other images apart from words. The origin of this communication comes from utilizing the resources available in typing, which is also known as ASCII art [84]. ASCII art is the ability to create an image based on the characters available on a typical keyboard. Some of the most common ASCII symbols are shown below in the table:

Words	**ASCII art**
Happy	:)
Sad	:(

Crying	:'(
Heart	<3

Table 9: Examples from ASCII Art [84]

In 1999, a stylized version of ASCII art was developed in Japan to convey information in a simple yet visually appealing way. Shigetaka Kurita created the original 176 images termed "emoji," which allow more emotion to be included in the communication [85].

5.3.2.6 Evaluating Written Communication

Being able to evaluate written communication is one of the most important tools in the latest industrial revolution. With countless amounts of information generated every day, the ability to process information in a timely fashion is of great value in making decisions quickly and properly. IBM has developed an Unstructured Information Management Architecture to aid in the process of information evaluation as shown in the figure below [82].

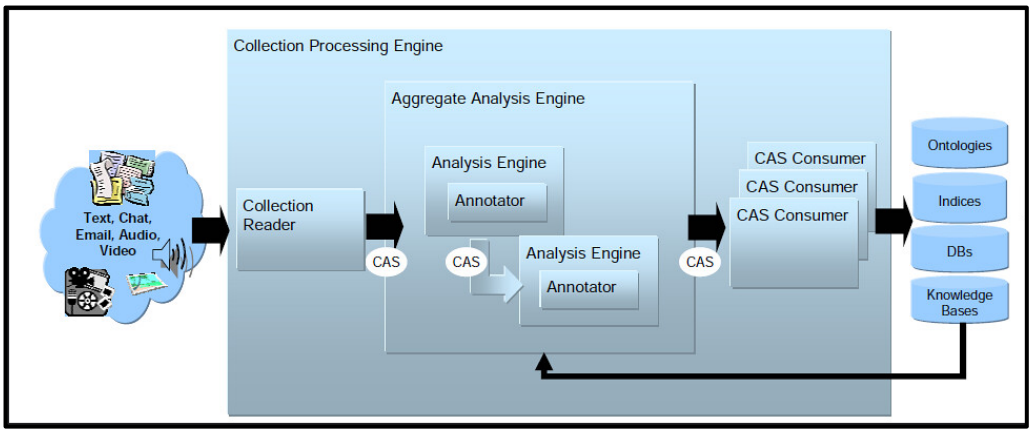

Figure 44: Unstructured Information Management Architecture [82]

The explanation of this process is given as follows: "Starting from the left, the Collection

Reader selects each document to process, and creates an empty Common Analysis Structure (CAS) which serves as a standardized container for information. This CAS is passed to Analysis Engines, composed of one or more Annotators which analyze the text and fill the CAS with the information found. The CAS are passed to CAS Consumers which do something with the information found, such as enter an entry into a database, update an index, or update a vote count" [82].

One large application area for text analysis is that of improving text search. By detecting important terms and topics within documents, semantic search engines provide the capability to search for concepts and relationships instead of just keywords as IBM's enterprise search solutions did [82]. In the model for writing correspondences, it is always tailored to the audience and how one wants the audience to react. In some cases, different audiences will react differently, so understanding the triggers of the audience would become the objective. When developing the content for a correspondence, there could be many different angles from which the document could be written. Some types of tone would be:

- Positive or negative
- Agressive, assertive, passive, passive agressive
- Leader, follower, alone, team member
- Empathetic, sociopathic, egotistical

- Professional, social, scholarly, common

To determine if a sentence matches a tone, it would have to be assessed on its patterns and anomalies. Patterns found within sentence structure are key to identifying the action, description and direction of the communication. If the pattern is recognized, then the elements of the pattern can reveal detailed information about the tone. The key words in the pattern, as well as the presence or lack of punctuation, would also serve as indicators of tone. In a simple case such as determining a positive or negative tone, these indicators would push and pull the intended tone. An example analyzing a single word in a phrase and its impact on tone is illustrated below:

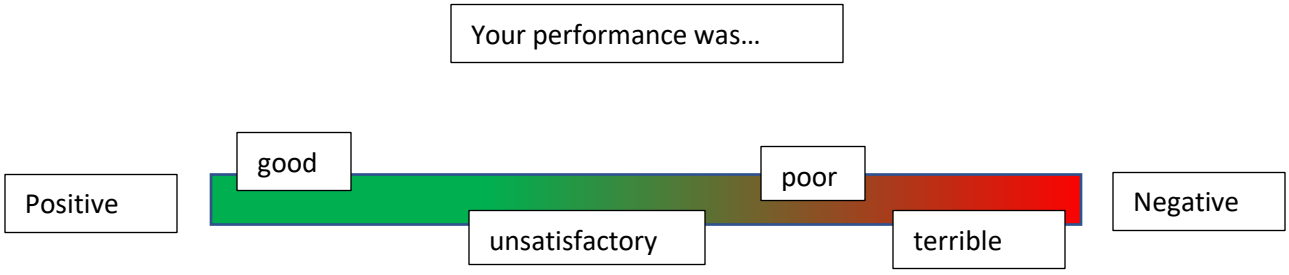

Figure 45: Analysis of Single Word in a Phrase

The degree to which each word causes a phrase to become more negative or positive is first subject to the rules imposed in the model. For different people, this model may have different weight factors and different responses.

A second type of scaling of a word could have multiple weights associated with a term depending on its context. Using an arbitrary scale of 1-10, each word could be identified by an

associated category, as shown in the figure below.

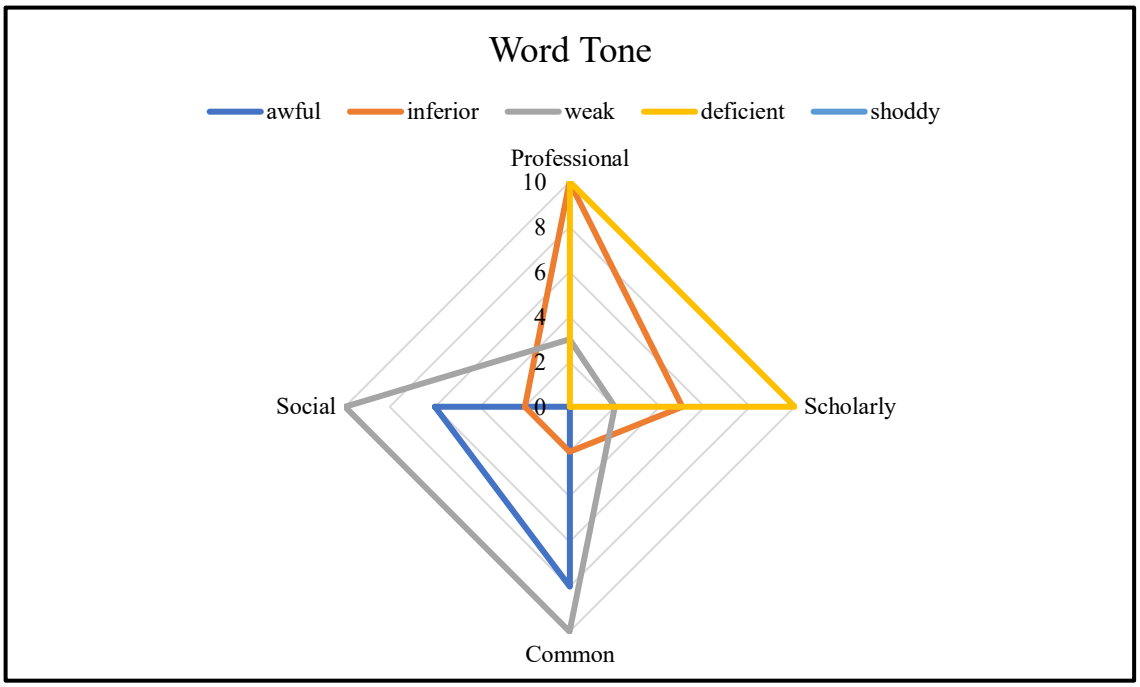

Figure 46: Word Tone

One of the key areas in digital technology research is the use of artificial intelligence to be able to draw correlations between words. Solutions like opencyc may be able to provide the engine for a computer to be able to understand the tone of a written correspondence [86]. The above-mentioned rewards function (equation 14) would then be a correlation model between the data and the results that are to be evaluated.

Depending on organizational needs, the rewards function can be modified to add more variables. If those variables have quantitative data, then a statistical model can be used to calculate the points; if not, then an unsupervised machine learning algorithm such as K-means

clustering, or NLP can be used to calculate the numerical values.

In the chosen points system, collaboration and efficiency have been selected as examples to show how points can be obtained. However, there are other ways that experienced employees can obtain further credits, such as participating in a mentoring program, choosing to go to additional training, and volunteering for community or company events.

5.4 Benefits of This Model

This model focuses on employees and would help tackle some of the challenges they face. Implementing this solution in organizations is beneficial for the following reasons:

- The new generation of employees is most likely stay with a company that recognizes their contributions, initiatives and gives them an opportunity to learn and grow [20].
- It helps show a company's vested interest in employees' wellbeing and social responsibility rather than just making a profit.
- It would increase productivity by being able to attract and retain talent.
- A data-based, subjective performance evaluation can be created and used to recognize top talent and promote them into leadership positions.

The model is employee-focused due to the reasons explained below:

- It advocates to give ownership of data to the employees and allow them to choose which data they share with the management team. This would also reduce liability on the part of

the organization.

- In this gamified system, a larger part of the rewards can be obtained by self-involvement and improvement rather than competing against other employees for rewards and points.
- Collaboration and cooperation are the main elements of the system that are being considered.
- Employees receive instant feedback to improve performance.
- This system gives employees the ability to track work progress and personal growth in the organization.
- The system can help employees set long-term and short-term goals.

5.5 Reward System

Current gamification methodologies use badges, points, and virtual and real prizes as rewards. In addition, a new reward system like those used in loyalty programs is selected here. There are two stages of the reward system:

- Acquiring reward points
- Spending rewards

The points could be earned through various work activities that may or may not be previously recognized. Such activities may include:

- Training new staff
- Shop safety and maintenance
- Exceptional customer satisfaction
- Working off-hours and in remote locations

Once earned, these points could go into various rewards including:

- Time off from work
- Opportunities to attend training courses
- Flexibility in terms of place and time of work
- Cash Bonus
- Task Auction

With each opportunity that can be purchased (spending rewards), it must be recognized that the cost for an opportunity would have to be balanced with the supply. Thus, people who perform tasks that may be undesirable to the group would earn higher points. The same would go for the purchase of the rewards as well.

The new generation of workers expects the above-mentioned flexibilities in a workplace. However, it might be unrealistic for an organization to offer this to everyone. Imagine no one shows up to work; instead, they choose to work from home, or everyone wants to go on a training

program and gets discouraged when they do not get an opportunity. In this methodology, it would be easier for companies to recognize the contributions employees are making even at a smaller scale, which would trigger intrinsic motivation and reward employees a productive way to trigger their extrinsic motivation.

6 Gamification and Law

Gamification is based on the fundamentals of behavioural theories and economics, as discussed in previous sections. In order to implement these theories, one needs to collect behavioural data from the users, which can then be analyzed using a rules engine. The gathered quantitative data can then be fed into a specific feedback mechanism to either help the user shape their behaviour or reward preferred behaviours through intrinsic or extrinsic motivation. Thus, it is important for developers to pay special attention to the laws that govern the collection of personal data.

6.1 Data Gathering

A successful gamification model is based on real-time data collected from users. Most of the gamified models that focus on shaping user behaviour would become irrelevant if the data collection were disabled for any reason. The type of data collected for gamified systems can include the following:

- GPS coordinates (location information)

- Time spent on a specific activity
- Time spent on a certain website
- Type of content browsed by an online user
- Social network data (comments, reviews, ratings, Facebook likes and dislikes, and Twitter feeds)

This data would then be used in gamified models as explained below. The following diagram shows the basic framework of gamification where big data is used in building the model [87].

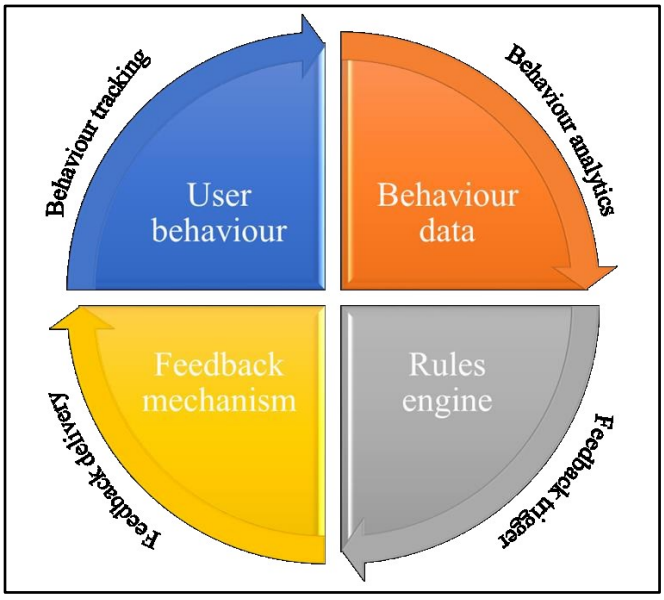

Figure 47: Basic Framework of Gamified System [87]

The user interface is the front port through which users interact with the gamified system.

This interface could be set up through websites, social media, or enterprise software. Behavioural data gathered through the system are used to create reports and dashboards to give a better understanding of users' behaviours. The analyzed raw data would then be fed into a rules engine, which decides the types of feedback to give the user based on its results. This rule engine is designed and configured by the gamified system developers and administrators. A feedback mechanism can then offer users rewards such as badges, points, and any other rewards built in by the administrator. The objective of any feedback system is to provide information to users that would help improve or modify their behaviour [87]. For example, consider a gamified exercise system that provides feedback to the user on how well they are doing, as shown in **Figure 28** [88]. If the focus of the user is to burn calories, then based on the feedback the user gains, he/she can set new goals to improve calorie burning if desired.

Figure 48: Example of Gamified Exercise System [88]

As the gamification model is built on gathering and using personal data, it is important to understand and comply with the laws that govern the collection and use of these data.

6.2 Legislative Requirements

Companies that use gamified systems should make sure that their policies to gather and use personal data are based on the relevant laws of the land where they are located. This section will focus on the federal privacy laws and regulations that govern this type of data collection in Canada.

6.2.1 Contract Law

In today's society, formal agreements are finalized using signed written contracts. These are legally binding and enforceable by law when signed by all the parties involved. These types of contracts are governed by civil law (in Quebec) or common law (everywhere else in Canada). In general, these two laws follow the same rules when it comes to contracts.

According to the law, for a contract to be legally binding, the following five conditions must be met [89].

1. Mutual free and informed consent from all the parties involved [89].
2. Contractual capacity, the mental capability to retain the agreement made; for example, someone who is underage (a child) or a person who is mentally ill would be unqualified to enter into a contract under this condition [89].

3. Contract should be for either goods or services [89].
4. The fourth condition is technically different in civil law and common law. In civil law, it requires a "lawful cause" to exist, whereas in common law, a "valuable consideration" is required. In brief, "…According to this fourth condition, the promise made must be serious and each obligation assumed by one of the parties must find a corresponding, but not necessarily equivalent or equal, promise made by the other party. A person may thus legally sell goods at a price that does not represent their actual market value. The contract would still be a valid one" [89].
5. Compliance with "formalities provided by law" [89].

The above conditions are just a brief description of the conditions that companies should focus on when drafting a contract that would allow users to consent to the gathering of their personal data. Without a proper contract in place, it is illegal to gather and use personal data. In addition to contract law, other types of laws, including copyright law, patent law and consumer protection law, should also be taken into consideration to protect both intellectual property and end users.

6.2.2 Case Study: Facebook Scandal

The Facebook - Cambridge Analytica scandal is a very recent controversy in this era of big data. Facebook, one of the largest social media and social networking platforms with over 2.13 billion monthly active users [90], is facing international investigations into allowing

Cambridge Analytica, a political consulting firm, to illegally gather and use personal data in order to shape public opinion. The details of the scandal are still being uncovered; however, the information already available in the public domain and the legislative changes being brought forward due to this scandal are vital to understand for anyone who plans to gather and use personal data for any purpose in the future.

Cambridge Analytica started gathering personal information from a small group of Facebook users in 2014 through a Facebook-linked app named "thisisyourdigitallife" developed by Aleksandr Kogan, a data scientist from Cambridge University. At the beginning, Kogan needed users' consent to gather data, and so provided money to Facebook users to take part in an elaborate personality test under the guise of academic research [91]. Those users consented to give their information, according to the Facebook Deputy General Counsel's initial statement: "The claim that this is a data breach is completely false. Aleksandr Kogan requested and gained access to information from users who chose to sign up to his app, and everyone involved gave their consent. People knowingly provided their information, no systems were infiltrated, and no passwords or sensitive pieces of information were stolen or hacked" [92]. However, instead of just gathering information from users who had signed up and given consent, the app also gathered data from other users in the friends lists of the willing participants. As per Facebook policy, this type of data harvesting is allowed only to improve the app's user experience and is prohibited from distribution for any other purposes, as stated on their website [2]. However, Cambridge

Analytica was able to use this method to harness data from 87 million Facebook users instead of just the few hundred thousand who did the initial survey. The collected data was then used "… to build psychographic profiles of users and their friends, which were used for targeted political ads in the UK's Brexit referendum campaign, as well as by Trump's team during the 2016 US election" [93]. This is considered the largest scandal involving misuse of personal data in the history of social media.

6.2.3 Legal Fallout

In a recent session, lawmakers in the United States Congress and Senate questioned the founder of Facebook, Mark Zuckerberg, on the mishandling of personal data. According to the *New York Times*, "…Mr. Zuckerberg seemed to suggest he could be open to some regulation, but neither he nor lawmakers seemed sure about how exactly to regulate the new breed of companies" [91]. In Europe, the European Union (EU)'s new privacy rules require Facebook to improve privacy safeguards. These European rules are designed to make the consent process less complicated. According to the EU, "… the rules are the most important change in data privacy regulation in a generation as it tries to catch up with technological advances since 1995, when the last comprehensive rules were approved" [94]. The EU rules also instruct that consent agreements must be written in a way that can be understood by anyone.

According to a recent Facebook blog post, "Everyone — no matter where they live — will be asked to review important information about how Facebook uses data and make choices

about their privacy on Facebook" [90]. The company further stated that they will begin implementing these choices first in Europe. As mentioned earlier, these legislative changes are being brought forward as recently as 2019. The laws governing the gathering personal data will continue to evolve and get tougher. Therefore, any developers who use personal data gathering should continue to pay close attention to these new developments and take the changes into consideration when building their models and getting appropriate user consent to obtain and utilize the data gathered.

7 Conclusion

The process of being able to turn big data into useful results requires as its backbone a well-developed model of how the data is to be collected and interpreted, as well as how the decisions from the results are to be implemented. Like all historical industrial revolutions, there will be a change in both the perceptions and the realities of how work is being done in the current industrial revolution, known as Industry 4. The key for success in implementing big data techniques is to fully develop the model for data collection, analysis, and decision-making. In this research we chose a model that has the potential to consider the impacts of data gathering on employees, society, and the environment; these factors may dictate the survivability of a company in terms of both productivity and societal perception. In data collection, being able to reference and cross-reference data using different perspectives is critical. In this study, a Weibull statistical analysis curve has been selected for evaluation of task completion duration, as it has a

simple dual variable system to define the shape and the stretch of the distribution. One of the shapes of this curve strongly resembles typical task completion prediction where, on occasion, task times are far longer than the average times. This method of tracking and analyzing data is beneficial to all parties involved. Therefore, it eases the reservations one might have in sharing personal data. With this, a gamified model is chosen to gamify employee behaviours that would benefit both organizations and employees.

The behaviours one wants to gamify are those that would optimize productivity. Productivity does not just depend on how much work is being done by whom, but also the quality of this work. It depends on how well the goals set by employees are aligned with company's short-term and long-term goals, as well as how closely the direction of an employee's work matches the company's strategic direction.

Currently, almost all human interaction in the workplace is competitive. The younger generation, however, does not like conflict at work. Conflicts arise among people when they are insecure about who takes credit for what. If the gamified model allows for the quantification of each worker's effort, rewarding them when they deserve credit, then that would reduce conflict and increase healthy competition. Some team-building behaviours can also be gamified to improve teamwork or collaboration through a rewards system. Such rewards could include extended vacations, time off from work, or credits towards attending company-sponsored courses.

One of the challenges faced when gamifying behaviour is collecting quantifiable data to form the basis for any payoffs. If there were no quantifiable data, then it would not be any different from traditional evaluation methods such as a manager or supervisor evaluating employees' performance based on personal perception. Traditional evaluation methods, though still used by many companies in Canada, are subjective and influenced by human emotions most of the time.

This research developed a methodology to gather and use quantifiable data to build a gamified model that rewards behaviours that help both the company and the employee grow. For this, a thorough study of motivation theories and general behaviour of the new generation of workers was included in this paper. The latter is required because as the generation changes, the needs and expectations of the workforce also change. For example, when the baby boomers occupied most of the workforce, their priorities were about stability; they were ready to work hard and long hours and wanted to stay at the same company. However, as the baby boomers are retiring, millennials are expected to replace them in large numbers, which will cause the overall expectations of the workforce to change. For these reasons, companies must look into innovative strategies such as gamification to keep the employees engaged and help them grow within the organization. Gamification can become an ideal tool in adjusting to these changing trends in the workforce.

References

[1] F. Oprescu, C. Jones, and M. Katsikitis, "I PLAY AT WORK-ten principles for transforming work processes through gamification," *Front. Psychol.*, vol. 5, 2014.

[2] "Facebook Platform Policy," *Facebook*. [Online]. Available: https://developers.facebook.com/policy. [Accessed: 19-Jul-2019].

[3] Prescient & Strategic Intelligence, "Gamification Market Overview." [Online]. Available: https://www.psmarketresearch.com/market-analysis/gamification-market. [Accessed: 20-Jul-2019].

[4] R. Garris, R. Ahlers, and J. E. Driskell, "Games, motivation, and learning: A research and practice model," *Simul. Gaming*, vol. 33, no. 4, pp. 441–467, 2002.

[5] A. Bandura, "Self-efficacy mechanism in human agency," *Am. Psychol.*, vol. 37, no. 2, pp. 122–147, 1982.

[6] A. H. A. H. H. Maslow, "A Theory of Human Motivation A Theory of Human Motivation," *Psychol. Rev.*, vol. 50, no. 4, pp. 370–396, 1943.

[7] A. C. Siang and R. K. Rao, "Theories of learning: A computer game perspective," in *Proceedings - IEEE 5th International Symposium on Multimedia Software Engineering, ISMSE 2003*, 2003.

[8] J. G. Nicholls, "Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance," *Psychol. Rev.*, vol. 91, no. 3, pp. 328–346, 1984.

[9] J. W. Atkinson and G. H. Litwin, "Achievement motive and test anxiety conceived as motive to approach success and motive to avoid failure," *J. Abnorm. Soc. Psychol.*, vol. 60, no. 1, pp. 52–63, 1960.

[10] K. Ling *et al.*, "Using social psychology to motivate contributions to online communities," *J. Comput. Commun.*, vol. 10, no. 4, 2005.

[11] K. Kraiger, J. K. Ford, and E. Salas, "Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation.," *J. Appl. Psychol.*, vol. 78, no. 2, pp. 311–328, 1993.

[12] T. Reiners and L. C. Wood, *Gamification in education and business*. Cham, Switzerland: Springer, 2015.

[13] K. Huotari and J. Hamari, "Defining gamification," in *Proceeding of the 16th International Academic MindTrek Conference on - MindTrek '12*, 2012.

[14] F. M. Ricciardi and E. Marting, *Top management Decision Simulation: the AMA Approach*. New York: American Management Association, 1957.

[15] K. J. Cohen and E. Rhenman, “The Role of Management Games in Education and Research,” *Manage. Sci.*, vol. 7, no. 2, pp. 101–194, 1961.

[16] A. J. Faria, D. Hutchinson, W. J. Wellington, and S. Gold, “Developments in business gaming: A review of the past 40 Years,” *Simul. Gaming*, vol. 40, no. 4, pp. 464–487, 2009.

[17] J. Wolfe, “A History of Business Teaching Games in English-Speaking and Post-Socialist Countries: The Origination and Diffusion of a Management Education and Development Technology,” *Simul. Gaming*, vol. 24, no. 4, p. 446, 1993.

[18] S. J. Schmidt, “Active and cooperative learning using web-based simulations,” *J. Econ. Educ.*, vol. 34, no. 2, pp. 151–168, 2003.

[19] P. Petridis *et al.*, “State-of-the-art in Business Games,” *Int. J. Serious Games*, vol. 2, no. 1, 2015.

[20] E. L. Deci and R. M. Ryan, *Handbook of self-determination research*. Rochester, NY: University of Rochester Press, 2004.

[21] B. Burke, “The Gamification of Business,” *Forbes Magazine*, 2013. [Online]. Available: https://www.forbes.com/sites/gartnergroup/2013/01/21/the-gamification-of-business/#2a7128ff4dc2.

[22] T. W. Kim and K. Werbach, “More than just a game: ethical issues in gamification,” *Ethics Inf. Technol.*, vol. 18, no. 2, pp. 157–173, 2016.

[23] S. Deterding, D. Dixon, R. Khaled, and L. E. Nacke, “From Game Design Elements to Gamefulness: Defining “Gamification",” in *MindTrek’11*, 2011.

[24] C. Malamed, “Are you a gamification wizard? Play the game.” [Online]. Available: http://theelearningcoach.com/reviews/gamification-of-learning/. [Accessed: 20-Jul-2019].

[25] “Leaderboard and Badging with Google Sheets.” [Online]. Available: http://www.teachingabovethetest.com/2017/12/leaderboard-and-badging-with-google.html. [Accessed: 20-Jul-2019].

[26] C. Alexander, I. Sara, and S. Murray, *A Pattern Language Towns, Buildings and Construction*. New York: Oxford University Press, 1977.

[27] E. Gamma, R. Helm, R. Johnson, and J. Vlissides, “Design Patterns: Abstraction and Reuse of Object-Oriented Design,” in *Software Pioneers*, New Dehli: Pearson India, 1994.

[28] J. Swacha and K. Muszyńska, “Design Patterns for Gamification of Work,” 2016.

[29] C. Zagel, *Service Fascination : Gaining Competitive Advantage through Experiential Self-Service Systems*. Wiesbaden: Springer Gabler.

[30] “Mathias De Brouwer, Sander De Wilde towards a healthy lifestyle MyPlan: A mobile application

and data analytics platform," Ghent University, 2016.

[31] M. Csikszentmihalyi, *Flow : The Psychology of Optimal Experience*. New York: Harper & Row, 1990.

[32] "Flow Theory." [Online]. Available: https://learn.canvas.net/courses/3/pages/level-6-dot-2-flow-theory. [Accessed: 20-Jul-2019].

[33] A. Pandey, "5 Killer Examples On How Gamification In The Workplace Is Reshaping Corporate Training," *eLearning Industry*, 2017. [Online]. Available: https://elearningindustry.com/gamification-in-the-workplace-reshaping-corporate-training-5-killer-examples. [Accessed: 20-Jul-2019].

[34] Melbournestorm, "Storm launches Australian first rewards program for members and fans." [Online]. Available: https://www.melbournestorm.com.au/news/2015/04/01/storm-launches-australian-first-rewards-program-for-members-and-fans/. [Accessed: 26-Jul-2019].

[35] Engine Yard, "Engine Yard: General." [Online]. Available: https://www.engineyard.com/faq. [Accessed: 26-Jul-2019].

[36] CallidusCloud, "Mysalesgame." [Online]. Available: https://www.calliduscloud.com/mysalesgame. [Accessed: 26-Jul-2019].

[37] Meeting Market, "Gamification for Motivation: Work Hard, Play Hard." [Online]. Available: https://meeting.market/gamification-for-motivation-work-hard-play-hard-gamification-for-motivation-work-hard-play-hard/. [Accessed: 26-Jul-2019].

[38] Bridge, "Case Studies." [Online]. Available: https://www.bridgesoftware.com/. [Accessed: 26-Jul-2019].

[39] Tovuti, "Tovuti." [Online]. Available: https://tovuti.io/. [Accessed: 26-Jul-2019].

[40] SABA, "Saba Cloud." [Online]. Available: https://www.saba.com/products. [Accessed: 26-Jul-2019].

[41] Skyprep, "All You Need to Optimize Your Online Training Platform." [Online]. Available: https://skyprep.com/. [Accessed: 26-Jul-2019].

[42] Docebo, "Docebo Learning Platform." [Online]. Available: https://www.docebo.com/. [Accessed: 26-Jul-2019].

[43] Ziiva, "Properity LMS." [Online]. Available: https://www.ziiva.com/. [Accessed: 26-Jul-2019].

[44] "Game Effective." [Online]. Available: www.gameffective.com. [Accessed: 01-Jul-2019].

[45] Y. Chou, "Octalysis – the complete Gamification Framework." [Online]. Available: https://yukaichou.com/gamification-examples/octalysis-complete-gamification-framework/.

[Accessed: 20-Jul-2019].

[46] M. Xu, J. M. David, and S. H. Kim, “The Fourth Industrial Revolution: Opportunities and Challenges,” *Int. J. Financ. Res.*, vol. 9, no. 2, pp. 90–95, 2018.

[47] O. of M. and B. Executive Office of the President, “North American Industry Classification System,” Washington,United States, 2017.

[48] Deloitte, “Challenges and Solutions for the Digital Transformation and Use of Exponential Technologies.” [Online]. Available: https://www2.deloitte.com/content/dam/Deloitte/ch/Documents/manufacturing/ch-en-manufacturing-industry-4-0-24102014.pdf. [Accessed: 20-Jul-2019].

[49] K. Schwab, *The Fourth Industrial Revolution*. London: Portfolio Penguin, 2017.

[50] P. N.Geoffrey and H. Aubin, “History of Europe,” *Encyclopædia Britannica, inc.*, 2019. [Online]. Available: https://www.britannica.com/topic/history-of-Europe/The-Industrial-Revolution.

[51] A. Wouter, S. Aaron De, and W. Kirsten, “Agility: It Rhymes with Stability,” *McKinsey & Company*. [Online]. Available: https://www.mckinsey.com/business-functions/organization/our-insights/agility-it-rhymes-with-stability. [Accessed: 20-Jul-2019].

[52] V. Lipman, “How To Manage Generational Differences In The Workplace,” *Forbes Magazine*, 2017. [Online]. Available: https://www.forbes.com/sites/victorlipman/2017/01/25/how-to-manage-generational-differences-in-the-workplace/#29983b714cc4. [Accessed: 20-Jul-2019].

[53] A. Khan and K. Turowski, “A Perspective on Industry 4.0: From Challenges to Opportunities in Production Systems,” in *International Conference on Internet of Things and Big Data*, 2016.

[54] Lmic, “What’s in a Name? Labour Shortages, Skills Shortages, and Skills Mismatches,” 2018. [Online]. Available: https://lmic-cimt.ca/wp-content/uploads/2018/10/LMI-Insights-No.-3.pdf. [Accessed: 20-Jul-2019].

[55] N. India - Surya, “The Skills Gap in U.S. Manufacturing,” *Deloitte*. [Online]. Available: http://www.themanufacturinginstitute.org/~/media/827DBC76533942679A15EF7067A704CD.ashx. [Accessed: 20-Jul-2019].

[56] R. Dewan, B. Jing, and A. Seidmann, “Product Customization and Price Competition on the Internet,” *Manage. Sci.*, vol. 49, no. 8, pp. 983–1119, 2003.

[57] A. Prossack, “How Millennials Are Changing The Way We View Leadership,” *Forbes Magazine*, 2018. [Online]. Available: https://www.forbes.com/sites/ashiraprossack1/2018/05/18/how-millennials-are-changing-the-way-we-view-leadership/#71004c415384. [Accessed: 20-Jul-2019].

[58] A. A. Facione, “Baby Boomers Retiring: Strategies for Small Businesses Retaining Explicit and Tacit Knowledge,” 2016.

[59] A. Alexandros and P. Dekoulou, *Cases on Corporate Social Responsibility and Contemporary Issues in Organizations*. Hershey, United States: IGI Global, 2019.

[60] Deloitte, "2018 Deloitte Millennial Survey Millennials Disappointed in Business, Unprepared for Industry 4.0." [Online]. Available: https://www2.deloitte.com/content/dam/Deloitte/global/Documents/About-Deloitte/gx-2018-millennial-survey-report.pdf. [Accessed: 20-Jul-2019].

[61] J. S. Kantor, "Want to Keep Your Millennials Mentor Them," 2016. [Online]. Available: https://www.huffpost.com/entry/give-them-what-they-want_b_8783712. [Accessed: 20-Jul-2019].

[62] J. Goudreau, "7 Surprising Ways To Motivate Millennial Workers," 2013. [Online]. Available: https://www.forbes.com/sites/jennagoudreau/2013/03/07/7-surprising-ways-to-motivate-millennial-workers/#3d070c11c79f. [Accessed: 20-Jul-2019].

[63] E. Garton, "Employee Burnout is a Problem with the Company, Not the Person." [Online]. Available: https://hbr.org/2017/04/employee-burnout-is-a-problem-with-the-company-not-the-person. [Accessed: 20-Jul-2019].

[64] V. Troaca, D. A. Bodislav, V. Troaca, and D. A. Bodislav, "Outsourcing. The Concept," *Theor. Appl. Econ.*, vol. XVIII(2012), no. 6(571), pp. 51–58, 2012.

[65] M. Abadijoo, M. N. Moghadam, and M. Beheshtifar, "Role of Human Resource Development in the Organization Success," *J. Soc. Issues Humanit.*, vol. 3, no. 7, pp. 44–47, 2015.

[66] C. E. Ebeling, *An Introduction to Reliability and Maintainability Engineering*. Long Grove: Waveland Press, 2010.

[67] M. Teimouri and A. K. Gupta, "On the Three-Parameter Weibull Distribution Shape Parameter Estimation," *J. Data Sci.*, vol. 11, pp. 403–414, 2013.

[68] C.-D. Lai, D. N. Murthy, and M. Xie, *Weibull Distributions and Their Applications*. Springer, 2006.

[69] N. R. Shankar, S. S. Babu, Y.L.P.Thorani, and D.Raghuram, "Right Skewed Distribution of Activity Times in PERT," *Int. J. Eng. Sci. Technol.*, vol. 3, no. 4, pp. 2932–2938, 2011.

[70] E. L. Mccombs, M. E. Elam, D. B. Pratt, E. L. ; Mccombs, and M. E. ; Elam, "Estimating Task Duration in PERT using the Weibull Probability Distribution," *J. Mod. Appl. Stat. Methods*, vol. 8, no. 1, pp. 282–288, 2009.

[71] J. Fente, K. Knutson, and C. Schexnayder, "Defining a Beta distribution function for construction simulation," in *WSC'99. 1999 Winter Simulation Conference Proceedings. "Simulation - A Bridge to the Future" (Cat. No.99CH37038)*, vol. 2, pp. 1010–1015.

[72] S. Ghosh and N. Das, "New Model of Performance Management and Measurement in Higher Education Sector," *Int. J. Sci. Res.*, vol. 2, no. 8, pp. 267–274, 2013.

[73] A. Shaout and M. K. Yousif, “Performance Evaluation – Methods and Techniques Survey,” *Int. J. Comput. Inf. Technol.*, vol. 03, no. 05, pp. 966–979, 2014.

[74] R. N. Lussier and J. R. Hendon, *Human Resource Management: Functions, Applications, Skill Development*. SAGE Publications, 2016.

[75] P. André, R. E. Kraut, and A. Kittur, “Effects of Simultaneous and Sequential Work structures on Distributed Collaborative Interdependent Tasks,” 2014.

[76] K. Zimmerman, “What Every Manager Should Know About Recognizing Their Millennial Employees,” *Forbes Magazine*, 2017. [Online]. Available: https://www.forbes.com/sites/kaytiezimmerman/2017/10/22/what-every-manager-should-know-about-recognizing-their-millennial-employees/#c7cb49a2d23b. [Accessed: 05-Jul-2019].

[77] E. Schreiner, “Importance of Cooperation in the Workplace,” 2016. [Online]. Available: https://woman.thenest.com/importance-cooperation-workplace-14647.html. [Accessed: 10-Jul-2019].

[78] J. P. Hespanha, *Noncooperative Game Theory An Introduction for Engineers and Computer Scientists*. Princeton, NJ: Princeton University Press, 2017.

[79] A. Kohll, “The Evolving Definition Of Work-Life Balance,” *Forbes Magazine*, 2018. [Online]. Available: https://www.forbes.com/sites/alankohll/2018/03/27/the-evolving-definition-of-work-life-balance/#57b32bfd9ed3. [Accessed: 15-Jun-2019].

[80] E. Liberty, “k-Means Clustering.” [Online]. Available: http://www.cs.yale.edu/homes/el327/datamining2013aFiles/10_k_means_clustering.pdf. [Accessed: 26-Jul-2019].

[81] E. Brill and R. J. Mooney, “Overview of Empirical Natural Language Processing,” *AI Mag.*, vol. 18, no. 4, pp. 13–24, 1997.

[82] IBM, “IBM Watson -- How to Replicate Watson Hardware and Systems Design for Your Own Use in Your Basement.” [Online]. Available: https://www.ibm.com/developerworks/mydeveloperworks/blogs/InsideSystemStorage/entry/ibm_watson_how_to_build_your_own_watson_jr_in_your_basement7?lang=en&CE=ISM0008. [Accessed: 27-Jul-2019].

[83] L. Felipe Gómez and S. L. Dailey, “Formal Communication,” in *The International Encyclopedia of Organizational Communication*, Hoboken, NJ, USA: John Wiley & Sons, Inc., 2017, pp. 1–15.

[84] “ASCII Art Archive.” [Online]. Available: https://www.asciiart.eu/. [Accessed: 29-Jul-2019].

[85] I. Moschini, “The ‘Face with Tears of Joy’ Emoji. A Socio-Semiotic and Multimodal Insight into a Japan-America Mash-Up,” *HERMES - J. Lang. Commun. Bus.*, no. 55, p. 11, Aug. 2016.

[86] A. Oltramari and C. Lebiere, “Extending cognitive architectures with semantic resources,” in

Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2011.

[87] M. Wu, "Is Gamification Dead?"

[88] "Where will Your Running Data be Tomorrow?" [Online]. Available: https://danaleeling.wordpress.com/2015/08/. [Accessed: 27-Jul-2019].

[89] J. Baudouin, "Contract Law in Canada," *The Canadian Encyclopedia*, 2016. [Online]. Available: https://www.thecanadianencyclopedia.ca/en/article/contract-law. [Accessed: 20-Jul-2019].

[90] E. Egan and A. Beringer, "Complying With New Privacy Laws and Offering New Privacy Protections to Everyone, No Matter Where You Live," *Facebook*, 2018. [Online]. Available: https://newsroom.fb.com/news/2018/04/new-privacy-protections/. [Accessed: 15-May-2019].

[91] N. Confessore, "Cambridge Analytica and Facebook: The Scandal and the Fallout So Far," *The New York Times*, New York, 04-Apr-2018.

[92] J. Winston, "The Facebook data breach is a scandal of our own making. Legally, there's nothing we can do about it.," 2018. [Online]. Available: https://www.nbcnews.com/think/opinion/facebook-data-breach-scandal- our-own-making-legally-there-s-ncna862211. [Accessed: 20-Jul-2019].

[93] I. Sherr, "Facebook, Cambridge Analytica and data mining: What you need to know," *Cnet*, 2018. [Online]. Available: https://www.cnet.com/news/facebook-cambridge-analytica-data-mining-and-trump-what-you-need-to-know/. [Accessed: 15-Jul-2019].

[94] "Facebook adds privacy settings to comply with new European law," *CBC News*, 2018. [Online]. Available: https://www.cbc.ca/news/technology/facebook-new-privacy-safeguards- europe-1.4624426. [Accessed: 20-Jun-2018].

Appendix A: IBM Watson

Role:	Do we need it for *personal use*?
Team Lead	Yes, That's you. Assuming this is a one-person project, you will act as Team Lead.
Algorithms	Yes, I hope you know computer programming!
Game Strategy	No, since *this version for personal use* won't be appearing on Jeopardy, we won't need strategy on wager amounts for the Daily Double, or what clues to pick next. Let's focus merely on a computer that can accept a question in text, and provide an answer back, in text.
Systems	Yes, this team focused on how to wire all the hardware together. We need to do that, although *this version for personal use* will have fewer components.
Speech Synthesis	Optional. For now, let's have *this version for personal use* just return its answer in plain text. Consider this *Extra Credit* after you get the rest of the system working. Consider using [eSpeak], [FreeTTS], or the Modular Architecture for Research on speech sYnthesis [MARY] Text-to-Speech synthesizers.
Annotations	Yes, I will explain what this is, and why you need it.
Information Sources	Yes, we will need to get information for *personal use* to process
Question Parsing	Yes, this team developed a system for parsing the question being asked, and to attach meaning to the different words involved.
Search Optimization	No, this team focused on making IBM Watson optimized to answer in 3 seconds or less. We can accept a slower response, so we can skip this.

Project Management	Yes, even for a one-person project, having a little "project management" never hurt anyone. I highly recommend the book [Getting Things Done: The Art of Stress-Free Productivity] by David Allen].

Appendix B: Weibull Distribution Equations

The calculation of the ratio

$$\frac{(M - x_A)}{(x_B - x_A)}$$

$$\frac{M - x_A}{x_B - x_A} = \frac{x_0 + \theta(1 - 1/\beta)^{1/\beta} - \left\{x_0 + \theta\left[\ln(1/R(x_a))\right]^{1/\beta}\right\}}{x_0 + \theta\left[\ln(1/R(x_b))\right]^{1/\beta} - \left\{x_0 + \theta\left[\ln(1/R(x_a))\right]^{1/\beta}\right\}}$$

Resulting Equation is:

$$\frac{M - x_A}{x_B - x_A} = \frac{(1 - 1/\beta)^{1/\beta} - \left[\ln(1/R(x_a))\right]^{1/\beta}}{\left[\ln(1/R(x_b))\right]^{1/\beta} - \left[\ln(1/R(x_a))\right]^{1/\beta}}$$

Table 1: Results of Fitting the Presented Weibull Model to the Data in Fente, et al. (2000)

Parameter	Eq.	Substituted Values	Result
β	(16)	$M = 9.21$, $x_a = 7.67$ $x_b = 11.05$, $R(x_a) = 1\text{-}0.01$ $R(x_b) = 1\text{-}0.75$	1.6900
$x_{a\ temp}$	(6)	$\theta = 1$, $R(x_a) = 1\text{-}0.01$, $\beta = 1.6900$	0.0657
$x_{b\ temp}$	(7)	$\theta = 1$, $R(x_b) = 1\text{-}0.75$, $\beta = 1.6900$	1.2132
σ^2_{temp}	(17)	$\theta = 1$, $\beta = 1.6900$	0.2953
K	(18)	$x_{b\ temp} = 1.2132$, $x_{a\ temp} = 0.0657$ $\sigma^2_{temp} = 0.2953$	2.1118
σ^2	(19)	$x_b = 11.05$, $x_a = 7.67$, $K = 2.1118$	2.5618
θ	(20)	$\sigma^2 = 2.5618$, $\beta = 1.6900$	2.9456
x_0	(21)	$M = 9.21$, $\theta = 2.9456$, $\beta = 1.6900$	7.4764
μ	(5)	$x_0 = 7.4764$, $\theta = 2.9456$, $\beta = 1.6900$	10.1056

Appendix C: Python Codes

Mapping of Table that contains Tracked data in MySQL

```
from sqlalchemy import Table, Column, DateTime, Integer, String, Text, \
   Date, Boolean, Float, DECIMAL, Numeric, BLOB, MetaData, ForeignKey, create_engine
from sqlalchemy.orm import sessionmaker
from sqlalchemy.ext.declarative import declarative_base

from dbMain import TrackingBase
engine = create_engine("mysql+pymysql://root:[redacted]@127.0.0.1:3306/project")
Base = declarative_base()

class ProgramLog(Base):
   __tablename__ = 'programlog'
   id = Column(Integer, primary_key=True)
   GUID = Column(String(255))
   Program = Column(String(255))
```

```
    WindowTitle = Column(String(500))
    Window2 = Column(String(500))
    WindowName2 = Column(String(500))
    TimeStart = Column(DateTime)
    KEvents = Column(Integer)
    MEvents = Column(Integer)
    SEvents = Column(Integer)
    TEvents = Column(Integer)
    Activity = Column(Integer)
    FDuration = Column(Integer)
    Tags = Column(String(255))

Base.metadata.create_all(engine)

Session = sessionmaker(bind=engine)
session = Session()

from sqlalchemy import Table, Column,DateTime, Integer, String, Text, Date, Boolean, Numeric,
MetaData, ForeignKey, create_engine
from sqlalchemy.orm import sessionmaker, relationship, Session, scoped_session, backref
from sqlalchemy.ext.declarative import declarative_base, declared_attr
import datetime
#from qtalchemy import *

class Base(object):
    mysql_engine='InnoDB',
    metadata = MetaData()


class TrackingBase(Base):
    __abstract__ = True
    __table_args__ = {'extend_existing':True}
    metadata = MetaData()
```

Table Mapper for Solidwors data table

```
from sqlalchemy import create_engine,Column, DateTime, Integer, String,func
from sqlalchemy.orm import sessionmaker
from sqlalchemy.ext.declarative import declarative_base
```

```
engine = create_engine("mysql+pymysql://root:[redacted]@127.0.0.1:3306/project")
Base = declarative_base()

class Solidworks(Base):
   __tablename__ = 'solidworks'
   Id = Column(Integer, primary_key=True,autoincrement=True)
   Features = Column(String(255))

   def __str__(self):
      return self.Features


Base.metadata.create_all(engine)

Session = sessionmaker(bind=engine)
session = Session()
```

Average Time Calculations

```
from Model import*
from solidbase import*
from dbTimeTracking import*

title=([])
names=[]
total=[]
key_words=[]
total=[]
nam=[]
sketchs=[]

extrude = []
sweep = []
fillet = []
round = []
thread = []
pattern = []
boss = []
```

```
fill = []
axis = []
sheet = []
weldment = []
trim = []
Three_3D = []
Cut = []
mirror = []
tube = []
HSS = []
revolve = []
split = []
loft = []
chamfer = []
structural_member = []
center = []
coordinate = []
rectangle = []
summation=[]
words= ['sketch', 'rectangle', 'parallelogram', 'slot', 'polygon', 'circle', 'arc', 'ellipse', 'parabola',
    'conic', 'spline', 'curve', 'extrude', 'revolve', 'sweep', 'boundary', 'fillet', 'round', 'chamfer',
    'thread', 'draft', 'rib', 'scale', 'dome', 'freeform', 'wrap', 'intersect',
    'pattern', 'mirror', 'loft', 'boss', 'hook', 'groove', 'vent','planar', 'flatten', 'offset',
    'radiate', 'fill', 'knit', 'split', 'composite', 'spiral', 'helix', 'axis', 'coordinate',
    'center', 'sheet',  'weldment', 'structural member', 'trim',
     'bead', 'tube','3D','Cut','HSS'] # key words that would be searched together with file names


Files=['AE 0217077 FIP','AE 0217077 MIP','AE 0217077  pup', 'AE 0217077 slip on flange', 'AE 0217077
Bottom Connector', 'AE 0217077 Hle','AE 0217077 Inlet down spout', 'AE 0217077 Ladder Base', 'AE
0217077 Ladder Link', 'AE 0217077 Ladder step', 'AE 0217077 Ladder Support', 'AE 0217077 Ladder', 'AE
0217077 Ladder2', 'AE 0217077 Lifting Lug', 'AE 0217077 Manhole Cover', 'AE 0217077 Manhole flange',
'AE 0217077 Manhole lid', 'AE 0217077 Manhole stoff', 'AE 0217077 Radiused wall', 'AE 0217077 Repad
Nozzle Bend', 'AE 0217077 Repad Nozzle', 'AE 0217077 Repad', 'AE 0217077 Repad2', 'AE 0217077
Repad3', 'AE 0217077 Repad5', 'AE 0217077 The base half', 'AE 0217077 Tank  half', 'AE 0217077
Reinforcement', 'AE 0217062 1 3050_BASKET Hrs bar stock   HRS', 'AE 0217062 1 3050_BASKET', 'AE
0217062 1 3050_BASKET ', 'AE 0217062 1 3050_BASKET ', 'AE 0217062 1 3050_BASKET', 'AE 0217062 1
3050_BASKET ', 'AE 0217062 1 3050_BASKET', 'AE 0218127 Beam Analysis', 'AE 0218127 Beam Analysis
With Plate', 'AE 0218127 D Anchor', 'AE 0218127 End Cap', 'AE 0218127 Post', 'AE 0218127 Post2',
```

'AE 0218127 Roof', 'AE 0218127 Steel Deck', 'AE 0218127 Zorbit', 'AE 0218127 Base Layer', 'AE 0218127 Beam Analysis ', 'AE 0218127 Beam Analysis With Plate ', 'AE 0218127 Post2 ', 'AE 0218041 Spacer', 'AE 0218041 Base Plate', 'AE 0218041 Beams', 'AE 0218041 Bench', 'AE 0218041 Bench2', 'AE 0218041 Bench3', 'AE 0218041 Bench4', 'AE 0218041 Bench5', 'AE 0218041 Channel', 'AE 0218041 Columns', 'AE 0218041 EndCap', 'AE 0218041 Gusset', 'AE 0218041 Lug', 'AE 0218041 Mount Plate', 'AE 0218041 Portal Crane Operator Cab', 'AE 0218041 Seat', 'AE 0218041 SeatBelt Anchor', 'AE 0218041 SeatBelt Anchor2', 'AE 0218041 SeatBelt Anchor3', 'AE 0218041 SeatBelt Anchor4', 'AE 0218041 SeatBelt Anchor1', 'AE 0218041 Spacer', 'AE 0218041 Plate', 'AE 0218041 Part1', 'AE 0218041 Back Rest', 'AE 0218041 Chair', 'AE 0218041 Beams ', 'AE 0218041 Bench ', 'AE 0218041 Bench2 ', 'AE 0218041 Bench3 ', 'AE 0218041 Bench5 ', 'AE 0218041 Columns Al 1.5 SQ X 0.1875', 'AE 0218041 Portal Crane Operator Cab ', 'AE 0218041 Portal Crane Operator Cab ', 'AE 0218041 Spacer', 'AE 0218041 Spacer ', 'AE 0218041 Spacer ','AE 0218041 Spacer ','AE 0218041 Spacer ','AE 0218041 Spacer ','AE 0218041 Spacer ','AE 0218041 Spacer ','AE 0218041 Spacer ','AE 0218041 Spacer ','AE 0218041 Spacer ','AE 0218041 Part1 ', 'AE 0218041 Back Rest ','AE 0217087 cra1','AE 0217087 Ferrarri Crane Base','AE 0217087 Knuckle','AE 0217087 Telescoping Boom Base','AE 0217087 Test Geometry','AE 0217087 Butt Section','AE 0217087 AdapterPlate','AE 0217087 Test Geometry ','AE 0217064 Base Plate','AE 0217064 FRAME','AE 0217064 Base','AE 0217064 Contact','AE 0217064 Frame 1','AE 0217064 Frame 2','AE 0217064 Hle','AE 0217064 Hitch Pin','AE 0217064 Proposed Extension','AE 0217064 SpreaderBar','AE 0217064 TEST1','AE 0217064 Wire Rope Guard','AE 0217064 FRAME ','AE 0217064 Frame 1 ','AE 0217064 Frame 2 ','AE 0217064 Proposed Extension ','AE 0218229 Frame','AE 0218229 LowerFrame','AE 0218229 Frame ','AE 0218197 Truck Frame','AE 0218197 Truck Frame ','AE 0218185 Pin','AE 0218175 Boom','AE 0218175 Winch Mount','AE 0218167 Liftin Beam','AE 0218167 Lug','AE 0218167 Pipe','AE 0218167 Plate','AE 0218167 Reel Guard','AE 0218167 Reel Guard2','AE 0218046 Pier','AE 0218046 End Beam','AE 0218046 Diagonal', 'AE 0218046 Compression Beam','AE 0218046 Bridge2','AE 0218046 Bridge','AE 0218043 Weld','AE 0218043 Weld2',
'AE 0218043 Fillet Weld','AE 0218043 Access Hatch','AE 0218043 Access HatchLatch1','AE 0218043 Access HatchLatch2','AE 0218043 Latch BasePlate','AE 0218043 Latch BasePlate2','AE 0218043 LatchPinHole','AE 0218043 Pin','AE 0218043 Pipe','AE 0218043 Sper','AE 0218043 Sper2','AE 0218043 Weld','AE 0218043 Weld2','AE 0218043 Weld ','AE 0218043 Weld ','AE 0218043 Weld ','AE 0218043 Weld ','AE 0218043 Weld ',
'AE 0218043 Weld ','AE 0218043 Weld ','AE 0218043 Weld ','AE 0218043 Weld ','AE 0218043 Weld2 ','AE 0218043 Weld2 ','AE 0218043 Weld2 ','AE 0218043 Weld2 ',
'AE 0218043 Weld2 ','AE 0218043 Weld2 ','AE 0218043 Weld2 ','AE 0218043 Weld2 ','AE 0218043 Weld2 ','AE 0218043 Weld2 ','AE 0218043 Weld2 ','AE 0217095 Bridge Model','AE 0217095 Embutmet','AE 0217095 Floor','AE 0217095 Girder Extension',
'AE 0217095 Girder Extension2','AE 0217095 Girder Removal Analysis','AE 0217095 Girder Removal Analysis2','AE 0217095 Girder Section','AE 0217095 Girder Section2','AE 0217095 Guard Rail','AE 0217095 L','AE 0217095 Pier2','AE 0217095 Rebar1','AE 0217095 RoadWay1','AE 0217095 RoadWay2','AE 0217095 Girder Removal Analysis ','AE 0218174 Boom Beam','AE 0218174 Jib Boom b',

'AE 0218174 Jib Boom','AE 0218174 Jib Boom O','AE 0218108 Frame','AE 0218108 Frame ','AE 0218108 Frame ','AE 0218015 Extension','AE 0218015 Proposed Extension','AE 0218015 Reel Evaluation', 'AE 0218015 Skid FEA Analysis','AE 0218015 St FEA','AE 0218015 Extension ','AE 0218015 Proposed Extension ','AE 0218015 Skid FEA Analysis ','AE 0218015 Skid FEA Analysis ','AE 0218015 St FEA ', 'AE 0218013 Attachment Point','AE 0218013 B','AE 0218013 Vessel Analysis','AE 0217100 Side Boom', 'AE 0217100 Side Boom2','AE 0217100 Side Boom3','AE 0217100 Side Boom ','AE 0217100 Side Boom ', 'AE 0217100 Side Boom2 ','AE 0217100 Side Boom2 ','AE 0217100 Side Boom3 ','AE 0217100 Side Boom3 ','AE 0217088 Elevator Guard2','AE 0217088 Elevator Guard2b','AE 0217088 OrginalElevator Guard','AE 0217088 Elevator Guard2 ','AE 0217088 Elevator Guard2 ','AE 0217088 Elevator Guard2b ', 'AE 0217088 Elevator Guard2b ','AE 0217088 OrginalElevator Guard ', 'AE 0217088 OrginalElevator Guard ','AE 0217081 Equipment Frame', 'AE 0217081 Part1','AE 0217081 Equipment Frame ','AE 0217081 Equipment Frame ','AE 0217081 Part1 ','AE 0217080 Column', 'AE 0217080 Floor', 'AE 0217080 Column ', 'AE 0217080 Floor', 'AE 45 Floor_1','AE 45 Floor_2','AE 45 Floor_3','AE 45 Floor_4','AE 45 Floor_5', 'AE 0217080 Floor', 'AE 0217080 Floor', 'AE 0217080 Floor', 'AE 0217080 Floor', 'AE 0217060 Rooster Sheave Repair','AE 0217052 Stickemup','AE 0217052 Stickemup 15', 'AE 0217052 Stickemup 15 2','AE 0217052 Stickemup ','AE 0217052 Stickemup ','AE 0217052 Stickemup 15 ','AE 0217052 Stickemup 15 ','AE 0217052 Stickemup 15 2 ','AE 0217052 Stickemup 15 2 ','AE 0217051 IPaper A','AE 0217051 IPaper B','AE 0217051 Part1','AE 0217051 Reinforcing Lug','AE 0217051 Part1 ', 'AE 0217049 1250bargrating','AE 0217049 Bracket','AE 0217049 Column Loads','AE 0217049 Pit Analysis', 'AE 0217049 Pit Analysis2','AE 0217049 Pit Analysis3','AE 0217049 Frame', 'AE 0217049 Plate', 'AE 0217049 U Frame','AE 0217049 U Frame2','AE 0217049 Pit Analysis ','AE 0217049 Pit Analysis2 ', 'AE 0217048 Stickemup 12 2','AE 0217048 Stickemup 15','AE 0217048 Stickemup','AE 0217048 Part1', 'AE 0217048 Lug','AE 0217048 Stickemup 12 2 ','AE 0217048 Stickemup 12 2 ', 'AE 0217048 Stickemup 15 ','AE 0217048 Stickemup 15 ','AE 0217048 Stickemup ','AE 0217048 Stickemup ','AE 0217047 Jack Securement Plate','AE 0217047 Outrigger','AE 0217047 Rotating Frame','AE 0217047 Spacer1', 'AE 0217047 Spacer2','AE 0217047 Spacer','AE 0217047 Spreader','AE 0217047 Sprocket','AE 0217047 Bolted Connection Test','AE 0217047 Collar','AE 0217047 Drop Leg Jack','AE 0217047 Engine Mount Plate','AE 0217047 Engine','AE 0217047 Jack Foot','AE 0217047 Outrigger Extension','AE 0217047 Pin', 'AE 0217047 Pin2','AE 0217047 Spreader Attachment','AE 0217047 Outrigger ','AE 0217047 Spreader ', 'AE 0217047 Outrigger Extension ','AE 0217043 TR_Booster','AE 0217037 Stick','AE 0217037 Stringing Jib','AE 0217037 Mock Beam','AE 0217037 Stick','AE 0217037 Stringing Jib','AE 0217034 Stick ','AE 0217034 Stringing Jib ','AE 0217034 Mock Beam','AE 0217034 Stick','AE 0217034 Stringing Jib','AE 0217033 Arms','AE 0217033 Cradle','AE 0217033 Hinge','AE 0217033 Pipelaying', 'AE 0217033 Pipelaying Eng Lug2','AE 0217033 Pipelaying FEA','AE 0217033 Wheel','AE 0217028 Angle Iron Base','AE 0217028 CrossBeam','AE 0217028 CrossBrace1','AE 0217028 CrosssBrace2','AE 0217028 CrosssBrace3','AE 0217028 Hrail 4','AE 0217028 Lug','AE 0217028 Wash Ramp Base A','AE 0217028 Wash Ramp Base B','AE 0217028 CrossBeam ','AE 0217028 CrossBrace1 ','AE 0217028 CrosssBrace2 ',

'AE 0217028 CrosssBrace3 ','AE 0217028 Hrail 4 ','AE 0217028 Wash Ramp Base A ','AE 0217028 Wash Ramp Base A ','AE 0217028 Wash Ramp Base B ','AE 0217026 Weld Repair','AE 0217025 2','AE 0217025 Boom Hling Lug','AE 0217022 Guardrail Shop Rev1','AE 0217022 Hrail Shop Rev1','AE 0217022 HRail2', 'AE 0217022 HRail3','AE 0217022 Stairs2','AE 0217022 Flange','AE 0217021 Anchor Point1','AE 0217021 Anchor Point2','AE 0217021 Anchor Point3','AE 0217021 Mounting Beam', 'AE 0217021 Weld on Plate', 'AE 0217021 Anchor Point2 ','AE 0217021 Mounting Beam ','AE 0217012 Containment3','AE 0217012 Containment3','AE 0217012 Containment2','AE 0217012 Containment','AE 0217012 Lock','AE 0217012 Hle','AE 0217012 Roof Panel1','AE 0217012 Roof Panel2','AE 0217012 Second Lug','AE 0217012 Shell','AE 0217012 Lug','AE 0217012 Window','AE 0217012 Window Frame1', 'AE 0217012 Window Frame2','AE 0217012 Window Frame3','AE 0217012 Hinge Base','AE 0217012 Containment3 ','AE 0217012 Lug ','AE 0217005 Truck Frame','AE 0217005 Truck Frame ','AE 0216020 Floor','AE 0216020 Post','AE 0216020 Base Plate','AE 0216020 Grouting','AE 0216019 Frame Assembly', 'AE 0216019 Frame Extension Arm','AE 0216019 Frame','AE 0216019 End Caps','AE 0216019 Hitch Pin', 'AE 0216019 Keeper1','AE 0216019 Latch','AE 0216019 Pipe','AE 0216019 Bar End Cap','AE 0216019 Frame Assembly ','AE 0216019 Frame Extension Arm ','AE 0216019 Frame ','AE 0216018 Jib', 'AE 0216018 Jib inside','AE 0216018 Tip','AE 0216018 Upper Boom Section','AE 0216018 Anchor', 'AE 0216018 Tip Section3','AE 0216018 Tip Section2','AE 0216018 Tip Section','AE 0216018 Spine', 'AE 0216018 Jib Boom','AE 0216010 Beam','AE 0216010 Bumper','AE 0216010 Bumper4','AE 0216010 Beam ','AE 0216010 Bumper ','AE 0216010 Bumper ','AE 0216010 Bumper4 ','AE 0216010 Bumper4 ', 'AE 0216008 Tank Dolly','AE 0216004 Mast Bottome FEA','AE 0216004 MastFEA','AE 0216004 Mast FEA', 'AE 0216004 Winch','AE 0216004 Lifting Point','AE 0216004 Mast Bottome FEA ','AE 0216004 Mast Bottome FEA ','AE 0216004 MastFEA ','AE 0216004 MastFEA ','AE 0216004 Mast FEA ','AE 0216004 Mast FEA ','AE 0117131 Part1b','AE 0117131 Part1','AE 0117082 Bearing Plate','AE 0117082 Bottom Connector','AE 0117082 Cable','AE 0117082 Connector','AE 0117082 Crane','AE 0117082 CraneEvaluation Original Truss','AE 0117082 CraneEvaluation Original','AE 0117082 CraneEvaluation Modification 1','AE 0117082 CraneEvaluation Modification 2','AE 0117082 CraneEvaluation', 'AE 0117082 CraneEvaluation2','AE 0117082 CraneEvaluation Propersize','AE 0117082 Overall Crane', 'AE 0117082 Overall Crane2','AE 0117082 BasicPin','AE 0117082 Plate1','AE 0117082 Plate3', 'AE 0117082 Connector','AE 0117082 CraneEvaluation Original Truss ','AE 0117082 CraneEvaluation Original ','AE 0117082 CraneEvaluation Modification 1 ','AE 0117082 CraneEvaluation Modification 2 ', 'AE 0117082 CraneEvaluation ','AE 0117082 CraneEvaluation2 ', 'AE 0117082 CraneEvaluation Propersize', 'AE 0117082 Overall Crane ','AE 0117082 Overall Crane2 ','AE 0218288 Rack']

```
Files= list(dict.fromkeys(Files))


for name in Files:
   na=name.split()
```

```
    nam.append(' '.join(na[0:2]))


    del na[:2]
    search=' '.join(na)
    key_words.append(search)


nam = list(dict.fromkeys(nam))


for i in key_words:
    file = "%" + i + "%"

    for ans in words:
        word = "%" + ans + "%"

        part = session.query(Solidworks).filter(Solidworks.Features.like(file) &
Solidworks.Features.like(word)).all()

        count=(len(part))

        if ans == 'sketch':
            sketchs.append(count)
        if ans=='extrude':
            extrude.append(count)
        if ans == 'sweep':
            sweep.append(count)
        if ans == 'fillet':
            fillet.append(count)
        if ans == 'round':
            round.append(count)
        if ans == 'thread':
            thread.append(count)
        if ans == 'pattern':
            pattern.append(count)
        if ans == 'boss':
            boss.append(count)
```

```
if ans == 'fill':
   fill.append(count)
if ans == 'axis':
   axis.append(count)
if ans == 'sheet':
   sheet.append(count)
if ans == 'weldment':
   weldment.append(count)
if ans == 'trim':
   trim.append(count)
if ans == '3D':
   Three_3D.append(count)
if ans == 'Cut':
   Cut.append(count)
if ans == 'mirror':
   mirror.append(count)
if ans == 'tube':
   tube.append(count)
if ans == 'HSS':
   HSS.append(count)
if ans == 'revolve':
   revolve.append(count)
if ans == 'split':
   split.append(count)
if ans == 'loft':
   loft.append(count)
if ans == 'chamfer':
   chamfer.append(count)
if ans == 'structural member':
   structural_member.append(count)
if ans == 'center':
   center.append(count)
if ans == 'coordinate':
   coordinate.append(count)
if ans == 'rectangle':
   rectangle.append(count)
```

```
        if count>0:


            total.append(count)


            for h in Files:
                names.append(h)

#Total entity calculation
summation=[]
entities=[]


for ans in key_words:

    for wr in words:

        word = "%" + wr+"%"
        file = "%" + ans + "%"

        part = session.query(Solidworks).filter(Solidworks.Features.like(file)&
Solidworks.Features.like(word)).all()
        summation.append(len(part))



d=len(words)
a=0

for element in summation:

    entities.append((sum(summation[a:d])))
    a=d
    d+=len(words)

    if d>len(summation):
        break
```

```
#-----------------------------------------------------------------------------------------------

# total time spent on each project
amount = []
for m in key_words:
   total_time = "%" + m + "%"

   time_spent =
session.query(ProgramLog).filter(ProgramLog.WindowTitle.like(total_time)&ProgramLog.WindowTitle.li
ke('%SLDPRT%')).all()
   amount.append(len(time_spent))

#average time-----------------------------------------------------------------------

ave = 0
average_time = []
for enti in  entities:
   if enti > 0:
      time=(int(amount[ave] / enti))
      average_time.append(time)
      ave += 1
      if ave > len(summation):
         break
   else:
      time=0
      average_time.append(time)
      ave += 1
      if ave > len(summation):
         break

#-----------------------------------------------------------------------------------
# filling the table in mysql database
countt=0

while countt<len(sketchs):
```

```
   Titles=Model(Title=Files[countt],sketch=sketchs[countt],extrude= extrude[countt], sweep=
sweep[countt],
            fillet=fillet[countt],round=round[countt],thread= thread[countt], pattern= pattern[countt],
            boss= boss[countt], fill= fill[countt],axis=axis[countt],sheet= sheet[countt],
            weldment=weldment[countt],
trim=trim[countt],Three_3D=Three_3D[countt],Cut=Cut[countt],mirror= mirror[countt],
            tube= tube[countt], HSS= HSS[countt], revolve= revolve[countt],split= split[countt], loft=
loft[countt],
            chamfer= chamfer[countt],structural_member= structural_member[countt], center=
center[countt],
            coordinate= coordinate[countt],rectangle= rectangle[countt],Total_Time_Spent
=amount[countt],Total_Number_of_Instances=entities[countt],Average_Time_Spent=average_time[cou
ntt])

   session.add(Titles)

   session.commit()

   countt+=1
```

Mapper for Model

```
from sqlalchemy import Table, Column, DateTime, Integer, String, Text, \
   Date, Boolean, Float, DECIMAL, Numeric, BLOB, MetaData, ForeignKey, create_engine
from sqlalchemy.orm import sessionmaker
from sqlalchemy.ext.declarative import declarative_base

from dbMain import TrackingBase
engine = create_engine("mysql+pymysql://root:E1ephant!@127.0.0.1:3306/project")
Base = declarative_base()


class Model(Base):
   __tablename__ = 'model'
   Id = Column(Integer, primary_key=True,autoincrement=True)
   Title = Column(String(255))

   Total_Time_Spent = Column(Integer)
   Total_Number_of_Instances = Column(Integer)
```

```
    Average_Time_Spent = Column(Float(4))

    sketch = Column(Integer)
    extrude = Column(Integer)
    sweep = Column(Integer)
    fillet = Column(Integer)
    round = Column(Integer)
    thread = Column(Integer)
    pattern = Column(Integer)
    boss = Column(Integer)

    fill = Column(Integer)
    axis = Column(Integer)
    sheet = Column(Integer)
    weldment = Column(Integer)
    trim = Column(Integer)
    Three_3D = Column(Integer)
    Cut = Column(Integer)
    mirror = Column(Integer)
    tube = Column(Integer)
    HSS = Column(Integer)
    revolve = Column(Integer)
    split = Column(Integer)
    loft = Column(Integer)
    chamfer = Column(Integer)
    structural_member = Column(Integer)
    center = Column(Integer)
    coordinate = Column(Integer)
    rectangle = Column(Integer)

#Base.metadata.drop_all(engine)

Base.metadata.create_all(engine)

Session = sessionmaker(bind=engine)
session = Session()
```

Mapper for Calculation model

```
from sqlalchemy import Table, Column, DateTime, Integer, String, Text, \
```

```
    Date, Boolean, Float, DECIMAL, Numeric, BLOB, MetaData, ForeignKey, create_engine
from sqlalchemy.orm import sessionmaker
from sqlalchemy.ext.declarative import declarative_base

from dbMain import TrackingBase
engine = create_engine("mysql+pymysql://root:E1ephant!@127.0.0.1:3306/project")
Bases = declarative_base()


class Calculationmodel(Bases):
    __tablename__ = 'calculationmodel'
    Ids = Column(Integer, primary_key=True,autoincrement=True)
    Titles = Column(String(255))

    Total_Time = Column(Integer)
    Total_Insts = Column(Integer)


    sketches = Column(Integer)
    extrudes = Column(Integer)
    sweeps = Column(Integer)
    fillets = Column(Integer)
    rounds = Column(Integer)
    threads = Column(Integer)
    patterns = Column(Integer)
    bosses = Column(Integer)

    fills = Column(Integer)
    axi = Column(Integer)
    sheets = Column(Integer)
    weldments = Column(Integer)
    trims = Column(Integer)
    Three_3Ds = Column(Integer)
    Cuts = Column(Integer)
    mirrors = Column(Integer)
    tubes = Column(Integer)
    HSSes = Column(Integer)
    revolves = Column(Integer)
    splits = Column(Integer)
```

```
    lofts = Column(Integer)
    chamfers = Column(Integer)
    structural_members = Column(Integer)
    centers = Column(Integer)
    coordinates = Column(Integer)
    rectangles = Column(Integer)

#Base.metadata.drop_all(engine)

Bases.metadata.create_all(engine)

Session = sessionmaker(bind=engine)
session = Session()
```

Weibull Calucations

```
from Model import*
from calculationmodel import*
import weibfun

Total_Number=[]
Total_time=[]
sket=[]
extru = []
swe = []
fil = []
rou = []
thre = []
patt = []
bos = []
fils =[]
ax = []
she = []
wel = []
tri = []
Thr= []
Cu = []
mir= []
tu = []
HS= []
```

```
revo = []
spl = []
lof = []
cham = []
structu = []
cent= []
coord= []
recta= []

for row in session.query(Model, Model.Total_Number_of_Instances).all():
   Total_Number.append(row[-1])

for row in session.query(Model, Model.Total_Time_Spent).all():
   Total_time.append(row[-1])


for row in session.query(Model, Model.sketch).all():
   sket.append(row[-1])

for row in session.query(Model, Model.extrude).all():
   extru.append(row[-1])

for row in session.query(Model, Model.sweep).all():
   swe.append(row[-1])

for row in session.query(Model, Model.fillet).all():
   fil.append(row[-1])

for row in session.query(Model, Model.round).all():
   rou.append(row[-1])

for row in session.query(Model, Model.thread).all():
   thre.append(row[-1])

for row in session.query(Model, Model.pattern).all():
   patt.append(row[-1])

for row in session.query(Model, Model.boss).all():
   bos.append(row[-1])
```

```
for row in session.query(Model, Model.fill).all():
   fils.append(row[-1])

for row in session.query(Model, Model.axis).all():
   ax.append(row[-1])

for row in session.query(Model, Model.sheet).all():
   she.append(row[-1])

for row in session.query(Model, Model.weldment).all():
   wel.append(row[-1])

for row in session.query(Model, Model.trim).all():
   tri.append(row[-1])

for row in session.query(Model, Model.Three_3D).all():
   Thr.append(row[-1])

for row in session.query(Model, Model.Cut).all():
   Cu.append(row[-1])

for row in session.query(Model, Model.mirror).all():
   mir.append(row[-1])

for row in session.query(Model, Model.tube).all():
   tu.append(row[-1])

for row in session.query(Model, Model.HSS).all():
   HS.append(row[-1])

for row in session.query(Model, Model.revolve).all():
   revo.append(row[-1])

for row in session.query(Model, Model.split).all():
   spl.append(row[-1])

for row in session.query(Model, Model.loft).all():
   lof.append(row[-1])
```

```
for row in session.query(Model, Model.chamfer).all():
   cham.append(row[-1])

for row in session.query(Model, Model.structural_member).all():
   structu.append(row[-1])

for row in session.query(Model, Model.center).all():
   cent.append(row[-1])

for row in session.query(Model, Model.coordinate).all():
   coord.append(row[-1])

for row in session.query(Model, Model.rectangle).all():
   recta.append(row[-1])

#-----------------------weighted time average-----------------------------------------------------------------------
ave = 0
sketches=[]
for total in  Total_Number:
   if total > 0:

      time=((sket[ave] / total)*Total_time[ave])
      if time>0:
         sketches.append(time)
      ave += 1
      if ave > len(Total_Number):
         break
   else:

      ave += 1
      if ave > len(Total_Number):
         break

sketches.sort()

analysis = weibfun.Analysis(sketches[140:200], unit='second')

analysis.fit(method='mle')
```

```
print(f'Beta: {analysis.Beta: .02f}')
print(f'eta: {analysis.eta: .02f}')


analysis.probplot()



analysis.pdf()
analysis.sf()
analysis.hazard()
analysis.cdf()
```

Appendix D: VBA Code

```
Dim swApp As Object
Sub main()

Dim swApp As SldWorks.SldWorks
Set swApp = Application.SldWorks

Dim fso As Object
Set fso = CreateObject("Scripting.FileSystemObject")

Dim xlApp As Excel.Application
Dim xlbook As Excel.Workbook
Dim xlSheet As Excel.Worksheet
Dim xlfiletemp As String
Dim basepath As String
Dim FileName As String
Dim filepath As String
Dim BuildDirectory As String
Dim answer1 As Integer
Dim db As Database
Dim rsEmployee As Recordset
Dim Counter As Integer
Dim rsEquipment As Recordset
Dim rsEquipmentType As Recordset
```

```
Dim swModel As SldWorks.ModelDoc2
Dim swFeat As SldWorks.Feature
Dim swSubFeat As SldWorks.Feature
Dim swDispDim As SldWorks.DisplayDimension
Dim swDim As SldWorks.Dimension
Dim swAnn As SldWorks.Annotation
Dim fileExists As Boolean
Dim toDay As Date
Dim fileNum As Integer
Dim numConfigs As Integer
Dim I As Integer
Dim confNames As Variant
Dim blnFeatFound As Boolean
Dim StoreValue(6)

Set swApp = CreateObject("SldWorks.Application")
Set swModel = swApp.ActiveDoc
Set swFeat = swModel.FirstFeature


Set xlApp = CreateObject("Excel.Application")
xlfiletemp = "D:\Thesis\Macro\FM-032-01.xlsm"

Set xlbook = xlApp.Workbooks.Open(xlfiletemp)
Set xlSheet = xlbook.Sheets(1)

xlApp.Visible = True

'ProjectNumber = InputBox("Enter Project Number:", "Project Number",
swModelDoc.CustomInfo("Project"))
ProjectNumber = InputBox("Enter Project Number:", "Project Number", "AE-02-16")

'xlbook.Sheets(1).Range("ProjectNumber") = ProjectNumber

'Set date
toDay = Now
StartRow = 1
StartGroupRow = StartRow
PreviousFeature = ""
ItemCount = 1

Dim fnameLength
```

```
ReDim Dimension(3, 0)

'How many configurations do we have?
numConfigs = swModel.GetConfigurationCount()

'Store the names of the configurations
confNames = swModel.GetConfigurationNames()

Counter = 0
Do While Not swFeat Is Nothing
  'Use this flags for printing the feature and sub-feature only once
  blnFeatFound = False


  Set swSubFeat = swFeat.GetFirstSubFeature
  Do While Not swSubFeat Is Nothing

      Set swDispDim = swSubFeat.GetFirstDisplayDimension

    Do While Not swDispDim Is Nothing
      Set swDim = swDispDim.GetDimension
           If Not blnFeatFound Then

        blnFeatFound = True

       End If
       'Print the sub-feature dims


       a = Split(swDim.FullName, "@")
       xlbook.Sheets(2).Activate
       xlbook.Sheets("Feature Tree").Cells(1, 1) = a(2)

       If PreviousFeature <> swFeat.Name Then

          If StartGroupRow <> Counter + StartRow Then
             Range(StartGroupRow + 2 & ":" & Counter + StartRow).Rows.Group
             xlbook.Sheets("Feature Tree").Cells(StartGroupRow + 1, 14) = "=if(SUM(N" &
StartGroupRow + 2 & ":N" & Counter + StartRow & ") = 0," & Chr(34) & Chr(34) & ", SUM(N" &
StartGroupRow + 2 & ":N" & Counter + StartRow & "))"
             xlbook.Sheets("Feature Tree").Cells(StartGroupRow + 1, 15) = "=if(SUM(O" &
StartGroupRow + 2 & ":O" & Counter + StartRow & ") = 0," & Chr(34) & Chr(34) & ", SUM(O" &
```

```
StartGroupRow + 2 & ":O" & Counter + StartRow & "))"
            xlbook.Sheets("Feature Tree").Cells(StartGroupRow + 1, 16) = "=if(SUM(P" &
StartGroupRow + 2 & ":P" & Counter + StartRow & ") = 0," & Chr(34) & Chr(34) & ", SUM(P" &
StartGroupRow + 2 & ":P" & Counter + StartRow & "))"
            xlbook.Sheets("Feature Tree").Cells(StartGroupRow + 1, 17) = "=if(SUM(Q" &
StartGroupRow + 2 & ":Q" & Counter + StartRow & ") = 0," & Chr(34) & Chr(34) & ", SUM(Q" &
StartGroupRow + 2 & ":Q" & Counter + StartRow & "))"
            xlbook.Sheets("Feature Tree").Cells(StartGroupRow + 1, 18) = "=if(SUM(R" &
StartGroupRow + 2 & ":R" & Counter + StartRow & ") = 0," & Chr(34) & Chr(34) & ", SUM(R" &
StartGroupRow + 2 & ":R" & Counter + StartRow & "))"
            xlbook.Sheets("Feature Tree").Cells(StartGroupRow + 1, 19) = "=if(SUM(S" &
StartGroupRow + 2 & ":S" & Counter + StartRow & ") = 0," & Chr(34) & Chr(34) & ", SUM(S" &
StartGroupRow + 2 & ":S" & Counter + StartRow & "))"
            xlbook.Sheets("Feature Tree").Cells(StartGroupRow + 1, 20) = "=if(SUM(T" &
StartGroupRow + 2 & ":T" & Counter + StartRow & ") = 0," & Chr(34) & Chr(34) & ", SUM(T" &
StartGroupRow + 2 & ":T" & Counter + StartRow & "))"
            xlbook.Sheets("Feature Tree").Cells(StartGroupRow + 1, 21) = "=if(SUM(U" &
StartGroupRow + 2 & ":U" & Counter + StartRow & ") = 0," & Chr(34) & Chr(34) & ", SUM(U" &
StartGroupRow + 2 & ":U" & Counter + StartRow & "))"
            xlbook.Sheets("Feature Tree").Cells(StartGroupRow + 1, 22) = "=if(SUM(V" &
StartGroupRow + 2 & ":V" & Counter + StartRow & ") = 0," & Chr(34) & Chr(34) & ", SUM(V" &
StartGroupRow + 2 & ":V" & Counter + StartRow & "))"
            xlbook.Sheets("Feature Tree").Cells(StartGroupRow + 1, 23) = "=if(SUM(W" &
StartGroupRow + 2 & ":W" & Counter + StartRow & ") = 0," & Chr(34) & Chr(34) & ", SUM(W" &
StartGroupRow + 2 & ":W" & Counter + StartRow & "))"
            xlbook.Sheets("Feature Tree").Cells(StartGroupRow + 1, 24) = "=if(SUM(X" &
StartGroupRow + 2 & ":X" & Counter + StartRow & ") = 0," & Chr(34) & Chr(34) & ", SUM(X" &
StartGroupRow + 2 & ":X" & Counter + StartRow & "))"
            xlbook.Sheets("Feature Tree").Cells(StartGroupRow + 1, 25) = "=if(SUM(Y" &
StartGroupRow + 2 & ":Y" & Counter + StartRow & ") = 0," & Chr(34) & Chr(34) & ", SUM(Y" &
StartGroupRow + 2 & ":Y" & Counter + StartRow & "))"
            xlbook.Sheets("Feature Tree").Cells(StartGroupRow + 1, 26) = "=if(SUM(Z" &
StartGroupRow + 2 & ":Z" & Counter + StartRow & ") = 0," & Chr(34) & Chr(34) & ", SUM(Z" &
StartGroupRow + 2 & ":Z" & Counter + StartRow & "))"
        End If

        StartGroupRow = Counter + StartRow
        Counter = Counter + 1
        xlbook.Sheets("Feature Tree").Cells(Counter + StartRow, 4) = swFeat.Name
        xlbook.Sheets("Feature Tree").Cells(Counter + StartRow, 3) = ItemCount

        ItemCount = ItemCount + 1
```

```
            Counter = Counter + 1
            xlbook.Sheets("Feature Tree").Cells(Counter + StartRow, 2) = "Custom"
        Else
            'StartRow = StartRow + 1
        End If

        Counter = Counter + 1

        PreviousFeature = swFeat.Name

        xlbook.Sheets("Feature Tree").Cells(Counter + StartRow, 2) = a(1)
        xlbook.Sheets("Feature Tree").Cells(Counter + StartRow, 5) = swDim.Name

        Select Case swDim.Name
            Case "W", "w", "Width", "Width"
                xlbook.Sheets("Feature Tree").Cells(Counter + StartRow, 14) = "=I" & Counter +
StartRow
                xlbook.Sheets("Feature Tree").Cells(Counter + StartRow, 15) = "=L" & Counter +
StartRow

            Case "H", "h", "Height", "height", "Tall", "tall"
                xlbook.Sheets("Feature Tree").Cells(Counter + StartRow, 16) = "=I" & Counter +
StartRow
                xlbook.Sheets("Feature Tree").Cells(Counter + StartRow, 17) = "=L" & Counter +
StartRow

            Case "T", "Thickness", "THK", "Wall"
                xlbook.Sheets("Feature Tree").Cells(Counter + StartRow, 18) = "=I" & Counter +
StartRow
                xlbook.Sheets("Feature Tree").Cells(Counter + StartRow, 19) = "=L" & Counter +
StartRow

            Case "OD", "Out_dia"
                xlbook.Sheets("Feature Tree").Cells(Counter + StartRow, 20) = "=I" & Counter +
StartRow
                xlbook.Sheets("Feature Tree").Cells(Counter + StartRow, 21) = "=L" & Counter +
StartRow

            Case "ID", "In_dia"
                xlbook.Sheets("Feature Tree").Cells(Counter + StartRow, 22) = "=I" & Counter +
StartRow
                xlbook.Sheets("Feature Tree").Cells(Counter + StartRow, 23) = "=L" & Counter +
```

```
StartRow

            Case "L", "Length"
               xlbook.Sheets("Feature Tree").Cells(Counter + StartRow, 24) = "=I" & Counter +
StartRow
               xlbook.Sheets("Feature Tree").Cells(Counter + StartRow, 25) = "=L" & Counter +
StartRow

        End Select

        SIValue = Round(swDim.GetSystemValue2("") * 1000, 1)
        ImperialValue = Round(swDim.GetSystemValue2("") * 39.37, 3)
        xlbook.Sheets("Feature Tree").Cells(Counter + StartRow, 11) = SIValue
        xlbook.Sheets("Feature Tree").Cells(Counter + StartRow, 12) = ImperialValue

'      End If

      Set swDispDim = swSubFeat.GetNextDisplayDimension(swDispDim)
    Loop
    'Get the next sub-feature
    Set swSubFeat = swSubFeat.GetNextSubFeature
  Loop
  'Get the next feature
  Set swFeat = swFeat.GetNextFeature
Loop

If StartGroupRow <> Counter + StartRow Then
    Range(StartGroupRow + 2 & ":" & Counter + StartRow).Rows.Group
End If

Set Part = Nothing
Set swApp = Nothing


Dim filepath2 As String

        filepath2 = "D:\Thesis\Macro" & ProjectNumber '& "\Documents\"
        'Call CreateFolderPath(filepath2, False)
        FileName = ProjectNumber & ".xls"

        xlbook.SaveAs (FileName)
```