

# Feature Based Statistical Model of Employee Productivity with Real Time Checked Data

Jalaja Shanmugalingam<sup>1</sup>, David Lario<sup>2</sup>, Yongsheng Ma<sup>\*1</sup>

<sup>1</sup>Department of Mechanical Engineering, University of Alberta, Edmonton, Canada

<sup>2</sup>Enigma Design Solutions, Grand Prairie, Canada

[yongsheng.ma@ualberta.ca](mailto:yongsheng.ma@ualberta.ca)

**Abstract** - The COVID-19 pandemic has led to a decentralization of the workforce in many industries. Due to the stay-at-home orders to control the spread of the virus, many are working from home. Even though modern technological advancements have helped some companies adapt to this new norm, many others are still scrambling to find the best way to remotely manage employees and accommodate their needs. Our research shows that the current challenges organizations face in managing their human capital are like the ones they face due to workplace demographic changes. This study focuses on analyzing those challenges and how human competency can be unlocked and developed to encourage sustainable autonomous working in an office, at home, or during frequent traveling. This study investigates the challenges faced by both organizations and employees, and presents a new business model that helps with the sustainable use of human resources and improves employee efficiency.

**Keywords** - Sustainable autonomous work; Weibull distribution; PERT; Statistical performance modeling; real-time efficiency data; Manpower management

## I. INTRODUCTION

The coronavirus pandemic has led to a new professional climate where companies are forced to decentralize their workforce and adapt to a remote working structure. On the other hand, employees who did not look favorably at the advanced use of technological tools now have no choice but to embrace this new normal. This change came into effect when organizations and their employees were least prepared for it [1]. It has become an additional challenge for companies that were already facing barriers in managing the expectations of a multigenerational workforce. Through this research, it was found that the challenges faced by the new generation of workers and organizations due to generational differences are very similar to the ones that have stemmed from the pandemic-related decentralization of the workforce.

A recent study shows that 34% of jobs in the United States can be performed from home [2]. Therefore, by dealing with the challenges related to decentralization, companies would not only be able to address the pre-existing concerns of the latest generation of workers, but also be able to easily take advantage of remote working even after the pandemic as "... (It) can be of great advantage specifically to large corporations for it offers more flexible hours of managing productivity." [3].

In this study, a review of the challenges faced by organizations and employees is carried out, followed by a description of a methodology aimed at addressing some of these challenges through implementing mentorship software that uses real-time data and statistical analysis. Following that, a case study is discussed, using real-time productivity data collected through tracking software developed by the researchers of this study to prove the viability of our methodology. Then, statistical modeling is used to find and discuss solutions to the above-mentioned challenges. In cases where data was not logged for individual components of work time, such as the time to complete a subtask (unmeasured) within a general task (measured), some methods to infer these aspects of the results are suggested. To determine the statistical probability of a result, the Weibull statistical curve is used. This model serves two functions: firstly, as a self-mentor which allows employees to understand and visualize their performance level on an ongoing basis using real-time data; and secondly, as an efficiency tool that allows organizations to identify their employees' strengths and weaknesses. With this information, companies will be able to utilize human resources in the areas where they are best suited, as well as offer opportunities for individual employees to improve their performance.

## II. STUDY OF CHALLENGES FACED BY ORGANIZATIONS

This section discusses the challenges faced by current business models in sustainably developing human capital to stay competitive at the time of pandemic-driven decentralization and presents a mechanism to tackle them effectively.

The following are some of the challenges businesses face in adapting to a new global economy in which decentralization is an increasingly vital component.

- Widening skills gap [4].
- Increased demand for customization [5].
- Leadership challenges (difficulty understanding the expectations of the new generation of workers) [6].
- Increased social responsibility [7].

### III. REVIEW OF CHALLENGES FACED BY EMPLOYEES

As the new generation of workers becomes more prominent in the workforce, companies need to widen their understanding of intergenerational expectation shifts and challenges to navigate the complex economic landscape. The following are some of those challenges as identified by researchers:

- Lack of mentorship [8], [9].
- Low motivation [10].
- Difficulty with work-life balance [11], [12].

### IV. RESEARCH QUESTION

In the past, organizations have focused mainly on increasing their financial efficiency by implementing traditional practices intended to increase employee productivity. This has helped to improve organizations' financial outcomes in the short term. However, in the long run, it impacted them negatively, leading to decreased productivity caused by employee burnout and loss of interest. In some industries, the solution companies came up with to deal with skill and labor shortages was outsourcing, "...outside resourcing," meaning to get resources from the outside" [13]. However, is this a long-term solution? This might increase the market capital in the short-term and generate wealth, but it does not constitute a long-term solution for the survival and sustainable growth of an organization. If companies are not ready to invest in human capital locally, then skill development will not happen. In fact, the skill shortage would become an even bigger problem, because properly utilizing human resources is one of the biggest contributors to professional success, both quantitatively and qualitatively. If harnessed well, the creativity, efficiency, and potential of human beings can help avoid many new challenges that might arise in the workplace. [14].

The companies that have withstood the test of time are the ones that focused on long-term solutions to the challenges they and their societies faced, as opposed to concentrating on short-term financial gains. At this time of accelerated economic growth, organizations are faced with additional challenges related to the overuse of resources, pollution, and climate change. This paper explores a methodology using big data to enable companies to better allocate, develop, and retain human talent, which in turn will promote long-term, sustainable innovation in a decentralized global economy.

### V. 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 through increased engagement, empowerment, and skill development. The optimization process of a work environment is based on goals and parameters that define success or failure within that workplace. Establishing those goals and parameters requires a mechanism to gather, analyze, and interpret

quantifiable data. With improvements in artificial intelligence and machine learning, computers can be used to sort through records and interpret results that may not be apparent with manual evaluation alone.

The diagram below shows the process of developing a dynamic model for this purpose, one that constantly collects data and continues to shape its parameters based on the data collected. A detailed explanation follows in the sections below.

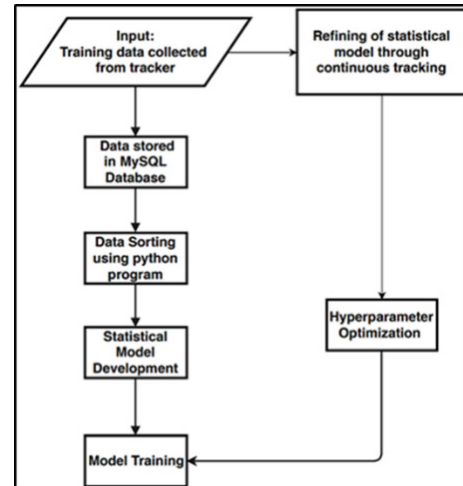


Fig. 1. Process of Model Development

The three core components of the proposed model are as follows:

- Data Collection
- Data Analysis
- Decision Making

#### A. Data Collection

In this research, the data was collected through a tracker program developed by Enigma Design Solutions that tracked every activity of its users. A detailed review of how this was conducted can be found in Section VII of this paper (titled "Case Study").

The main independent variable that is being 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, decision-making, productivity, setting measurable goals, creating learning opportunities, reducing stress, maintaining work-life balance, reducing financial losses, improving work quality, and establishing sustainable practices.

#### B. Data Sorting and Model Development

Data collected for this specific research is based on time spent on engineering projects and the different components of that project. For example, let's consider projects as a universal set and an engineering design as a subset with the following elements: modeling, analysis,

calculations, and sometimes research, as shown in the Venn diagram below.

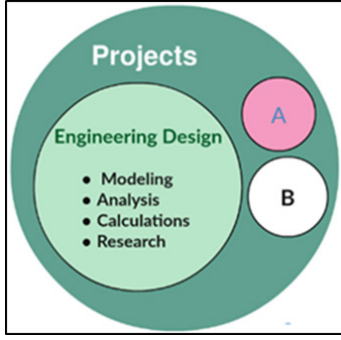


Fig. 2. Venn Diagram of Engineering Project

In a typical project like the one illustrated above, if we consider a single task in detail, such as the solid modeling of a single component, each part created will have different features and each person working on a part may use different methods to complete the task. In the past, there might have been no patterns or useful information that could be obtained from this drafting process. However, through the use of modern technological methods like data mining and artificial intelligence, complex patterns and habits can be extracted that would otherwise not be apparent.

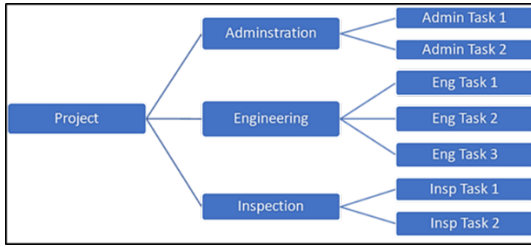


Fig. 3. Components of an Example Project

Let's consider solid modelling of a simple component in Solidworks. For this component, the geometry is made by creating sketches, which are in turn used to create features. Additional features can then be built on top of existing features. Using CAD software based on parametric modelling, each step is stored in a hierarchical format with a strong parent-child relationship that has sketches and features that are dependent on its parent. Any changes to the parent would have direct implications on the child features (i.e., changing geometry or removing the parent). The other feature of parametric modelling is that each sketch and feature is an object which has properties that can be extracted. Things like the number of types and quantities of a feature are easily determined. When extracting data on the use of such a program, 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 ( $t_n$ ).

$$\sum_n^m t_n = T \quad (1)$$

If the time for each feature cannot be measured, but the total time taken to create a model is known, then an assumption of "time to create" the feature could be used, where  $T$  is the total time, and  $A_n$  is the time-weighted average associated with it, calculated as shown below:

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

$$A_n = \frac{\text{Time taken to model all the same features}}{\text{Time taken to complete the model}} \quad (3)$$

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:

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

The average time taken to complete a feature can be estimated using projects as small samples. 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 a level of confidence. It would give the best estimate as to the average time one would take to complete a specific feature or task.

## VI. 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 [15].

### A. Weibull Distribution

Weibull Distribution, first introduced by physicist Waloddi Weibull, has the capability to model various types of distributions simply by changing parameters. The three-parameter Weibull can be described as follows.

Probability density function is given by:

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

The cumulative distribution function is given as:

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

Mean is calculated using:

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

Variance can be found by:

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

Where:

t: Random variable

$\theta$ : Scale parameter

$\beta$ : Shape parameter

$\gamma$ : Location parameter

$\Gamma$ : Gamma function

$t_0$ : Shifts the mean on the t axis

and:

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

### B. 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 [16].

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

$$-\frac{n\beta}{\theta} + \beta \sum_{i=1}^n \left(\frac{t_i - \gamma}{\theta}\right)^{\beta} = 0 \quad (2)$$

$$-(\beta - 1) \sum_{i=1}^n \frac{1}{t_i - \gamma} + \beta \sum_{i=1}^n \left(\frac{t_i - \gamma}{\theta}\right)^{\beta-1} = 0 \quad (3)$$

Where  $\beta > 0$ ,  $\theta > 0$ , and  $-\infty < \gamma < +\infty$

### C. Why Weibull Distribution?

Weibull distribution is capable of modelling symmetric data such as normal distribution, as well as 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  $\gamma$  is assumed to be zero.

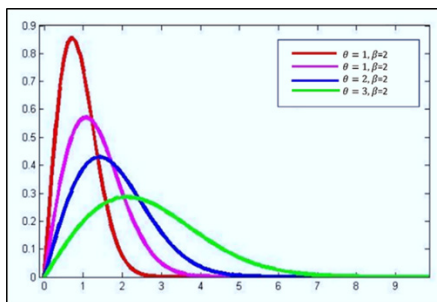


Fig. 4. Weibull Density Plots for Varying  $\beta$  and  $\theta$

To this day, popular Weibull distribution applications are found in reliability engineering, failure analysis, wind-

speed data analysis, unemployment duration, survival data [17], and PERT (Program Evaluation and Review Technique) [18]. The use of Weibull distribution in PERT is like the one discussed here. However, the application of improving sustainable productivity using big data is where this research differs from every other available application.

## VII. CASE STUDY

A tracker program was developed by Enigma Design Solutions in Python and installed on company computers to track user activity. 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;
- MEEvents: tracks the mouse usage (such as left and right clicks);
- SEEvents: 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 tallying the event count. This can then be utilized to calculate the time taken for each component of the project using a time-weighted average method, which can then be fed into a statistical model. The figure below shows the workflow of this process.

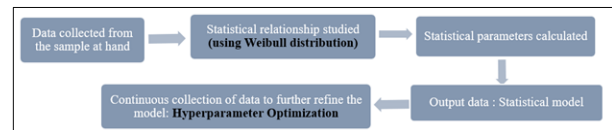


Fig. 8. Workflow of Model Development

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.

ID	Program	Window Title	Time	Seconds	Milliseconds	Activity	Duration	Tag
500123	python.exe	AE-02-08-03 - Web Way Access Modification	2018-04-02 21:41:41	0	2	0	2	1
500124	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:08	0	5	0	5	1
500125	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500126	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500127	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500128	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500129	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500130	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500131	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500132	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500133	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500134	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500135	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500136	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500137	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500138	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500139	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500140	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500141	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500142	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500143	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500144	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500145	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500146	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500147	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500148	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500149	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500150	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500151	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500152	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500153	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500154	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500155	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500156	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500157	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500158	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500159	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500160	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500161	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500162	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500163	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500164	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500165	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500166	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500167	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500168	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500169	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500170	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500171	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500172	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500173	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500174	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500175	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500176	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500177	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500178	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500179	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500180	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500181	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500182	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500183	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500184	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500185	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500186	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500187	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500188	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500189	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500190	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500191	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500192	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500193	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500194	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500195	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500196	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500197	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500198	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500199	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1
500200	python.exe	AE-02-08-03 - Web Way Access Modification	2018-05-02 13:08:13	0	4	0	4	1

Fig. 5. Collected Data

For this data to be associated with a specific project, it is important to title the window 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.

#### A. Data Sorting and Statistical Modelling

As mentioned earlier, the data collected is based on time spent on a specific project. For example, Figure 9 shows the time spent on a project that involves Solidworks modeling and analysis. This can be sorted further to study the time spent by the user on individual components of the larger task.

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

The following images show an example of data extracted via the above method. This can be done for every part that is modeled in Solidworks for a specific project.

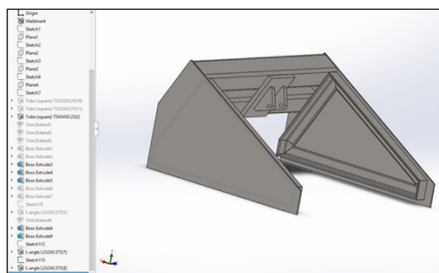


Fig. 6.VBA Macro for Solidworks Model

TABLE I  
EXTRACTED DATA

List of Features		
Comments	Front Plane	Design Binder
Favorites	Cut-Extrude3	Annotations
Equations	Cut-Extrude 1	Surface Bodies
Selection Sets	Sketch 2	Material

The collected Excel data is then transferred into a MySQL database through a program developed in SQLAlchemy for each model. The breakdown of the above information is then used as data for the time variable, which trains the statistical model and improves its future accuracy. The figure below shows how the Solidworks data is transferred into the MySQL database for each project.

Id	Features
52	AE-0218041-Beams Tube (square) TS1.5X1.5X0....
53	AE-0218041-Beams Tube (square) TS1.5X1.5X0....
54	AE-0218041-Beams Custom
55	AE-0218041-Beams Sketch13
56	AE-0218041-Beams Sketch13
57	AE-0218041-Beams Sketch13
58	AE-0218041-Beams Sketch13
59	AE-0218041-Beams Sketch14
60	AE-0218041-Beams Tube (square) TS1.5X1.5X0....

Fig. 7. Transferred Data

From the data extracted, the total number of instances for each feature can then be calculated, as shown in the figure below.

#	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	35	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	25	4	2	0	156	36	4
54	AE 0218041 Column	14	4	2	4	261	30	8
55	AE 0218041 EndCap	12	8	0	2	821	30	21

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 the sum of those features. This can then be used to calculate the average time taken to finish one individual feature in that set.

AE 0218041 Spacer: Feature Item Count

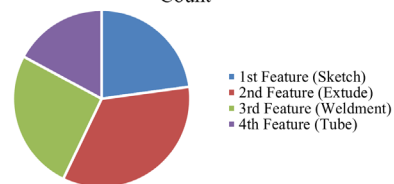


Fig. 9.Item Count

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 the 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 change as more data is added to the distribution. The addition of data changes the parameter values of beta and eta, as well as the overall shape of the distribution. A higher beta value leads to a normal distribution, whereas a beta value of less than one produces an exponential distribution, as shown below:

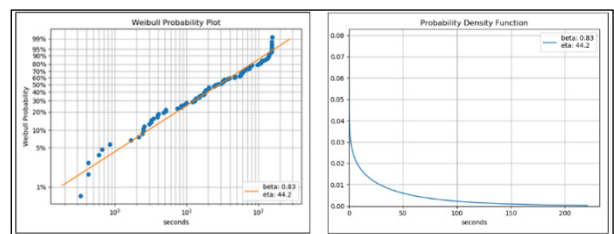


Fig. 10.First Weibull Probability Plot and Probability Density Function for Feature Sketch

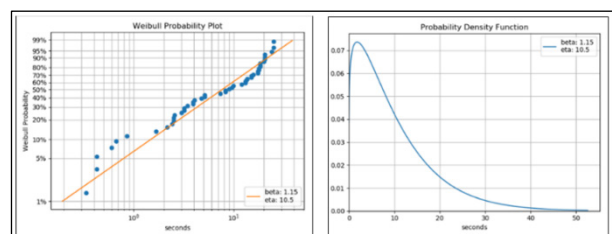


Fig. 11.Second Weibull Probability Plot and Probability Density Function for Feature Sketch

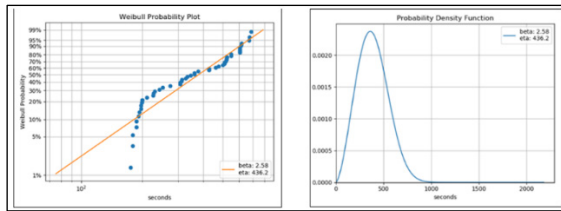


Fig. 12. Third Weibull Probability Plot and Probability Density Function for Feature Sketch

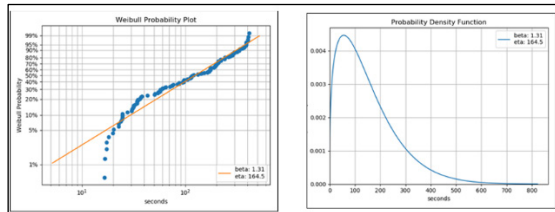


Fig. 13. First Weibull Probability Plot and Probability Density Function for Feature Extrude

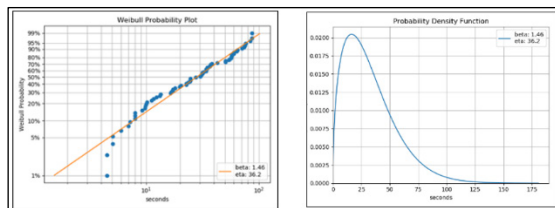


Fig. 14. Second Weibull Probability Plot and Probability Density Function for Feature Extrude

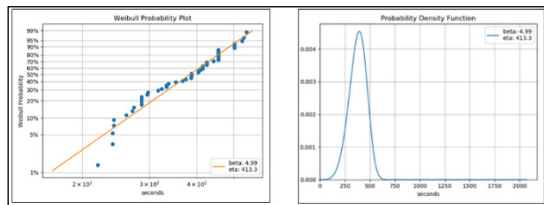


Fig. 15. Third Weibull Probability Plot and Probability Density Function for Feature Extrude

### B. Dependency and Concurrency

When considering this model, one must also consider the interrelationships between various tasks or features. Most of the tasks performed during work hours are often dependent on previous or concurrent tasks. An example of one such case is shown below, where the time taken for an intermediate step would depend on how quickly and how well the previous one was done.

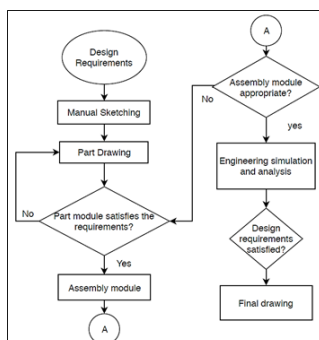


Fig. 16. Dependency Diagram

Instead of one exact value, this statistical model gives a range of possible values for a specific task's completion time. As mentioned earlier, sample averages are obtained to model the probability distribution, which is then used to estimate the population mean with a preferred confidence level for normally distributed data.

Once these data have been modeled, one can then 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 of tasks;
- The employee being in training, or;
- Interdependent tasks being delayed.

On the other hand, if the distribution is narrowly shaped, it would suggest the opposite of the above situations.

Based on this, one could analyze the factors influencing the completion of a specific task and how the completion time could be shortened. In the future, the tracker program could also be improved to track the time spent on not just the overall task but every intermediate task as well.

## VIII. DISCUSSION

Once the data are gathered and the statistical model implemented, what are their applications? As explained in the introductory paragraphs, every 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 issues that hinder sustainable innovation in various fields?

### A. Businesses

Optimization and human capital management: Human capital optimization is essential for businesses to increase their productivity and profit without sacrificing employee well-being or abandoning/avoiding the adoption of more sustainable business practices. Just as engineers choose materials and tools based on how well they will fit design requirements, knowing employees' strengths and weaknesses can help a company utilize the right talent in the right place to achieve overall goals without wasting resources. The knowledge needed to accomplish this can be gained using the above model.

The model would allow for an accurate task time estimate with a specific confidence level. Since the model is dynamic, companies can trust that any data gained would be current and applicable. This estimation can then be utilized to:

- Accurately estimate project completion time;
- Predict the exact cost of human capital, and;
- 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" [19]. Traditionally, performance management is done through performance appraisal, "...a formal system of review and evaluation of

individual or team task performance,” conducted 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 [20].
- Management by Objectives (MBO): Evaluated against goals set by management and achievement by the employee [20].
- Behaviorally Anchored Rating Scales (BARS): “BARS contrast an individual’s performance against specific examples of behavior that are anchored to numerical ratings” [20].
- Assessment Center: Behaviors are observed in set exercises as an evaluation method [20].

The above methods of performance evaluation are subjective and often not very positive. A recent survey shows that 80 percent of employees expressed displeasure with their performance evaluation system [21]. Performance evaluations are usually done for recruitment, training and career planning, for compensation programs and to assess the potential of an employee [21].

With the model explored in this study, instead of relying on subjective methods, companies can use hard data to build a system to evaluate an employee’s ability to perform tasks, as well as the actual contributions he or she makes to the company.

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

Offering 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 thus help companies draw and retain talent.

#### B. Employees

Self-evaluation: The above model could be used by employees on an individual basis so they can identify their strengths, project contributions, and potential areas of improvement. This self-evaluation would be purely based on data collected rather than subjective performance evaluation.

Virtual mentor: A mentor is someone who shares knowledge, gives advice based on career experience, and offers guidance when needed to help mentees succeed in their chosen professions. As millennials are taking over the workforce and learning the ways of industry, they need to have the right mentors to guide them. There can often be a gap between what industries require 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 help by acting as a virtual mentor, guiding employees to build their careers through an increased understanding of their capabilities and potential.

#### C. Disadvantages

One might ask why employees would willingly sign up to have Big Brother watch over their shoulders and allow the organization to scrutinize the way they spend their time at work. To encourage employees to take advantage of this method, companies can:

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

Employees are more inclined to accept changes and adapt when they see it benefits their growth and stability.

#### IX. FUTURE WORK & GENERALIZED MODEL

The base knowledge for this study was retrieved from research in multiple domains such as the study of workplace dynamics, demographics, and efficiency. Based on the results obtained, we concluded that an efficiency model using real-time data would help in resolving some of these problems by acting as both a management tool and a self-mentor as explained above. Likely, the above model can be generalized for various types of tasks and fields, not just engineering. Using new technology, activity tracking can be done for anyone, not just those who work on computers. For instance, there are devices that could collect data using other tracking methods (such as those used in this research), like GPS, the internet, or any other digital medium. Also, a mathematical optimization model can be developed to run in conjunction with the dynamic data model to aid in project management.

#### X. CONCLUSION

The dynamic data model discussed in this study has great potential and a wide variety of benefits it can bring to companies and workers in different fields. In this day and age, a company’s ability to meet challenges surrounding human capital management can often dictate that company’s survivability in terms of both productivity and societal perception. Each component of this model has its role to play in fostering more sustainable and productive business practices. Firstly, consistent, dynamic, and precise data collection, as well as the ability to cross-reference these data in many ways, allows work processes to be categorized and analyzed from different critical perspectives. This is essential for a company in the 21<sup>st</sup> century to develop plans to handle increased decentralization. Secondly, the use of the Weibull statistical analysis curve and its dual variable system brings an added layer of confidence to the data analysis process.

This confidence is often lacking in more subjective or traditional performance appraisal methods. Thirdly, the precision and adaptability of this tool will be essential for upper management in making broad decisions about how to best allocate company resources to utilize talent, boost productivity, and reduce wastage. Lastly, the personal ownership of data and possible use of this tool as a self-mentor will allow employees across industries to grow, improve, and further their career prospects on their own terms. Freedoms like these are very important to the new generation of workers and, since these individuals form the backbone of development and innovation, meeting these needs will be crucial to the sustainable construction of our collective future during this time of increased decentralization.

## REFERENCES

- [1] D. Savić, "COVID-19 and work from home: Digital transformation of the workforce," *Grey J.*, vol. 16, no. 2, pp. 101–104, 2020.
- [2] J. Dingel and B. Neiman, "How Many Jobs Can be Done at Home?," *Covid Econ.*, vol. 1, no. April, pp. 16–24, Apr. 2020.
- [3] Dr. Marivic F. Flores, "Understanding The Challenges Of Remote Working And It's Impact To Workers," *Int. J. Bus. Mark. Manag.*, vol. 4, no. 11, pp. 40–44, 2019.
- [4] N. India - Surya, "The Skills Gap in U.S. Manufacturing," *Deloitte*. [Online]. Available: <http://www.themanufacturinginstitute.org/~media/827DBC76533942679A15EF7067A704CD.ashx>. [Accessed: 20-Jul-2019].
- [5] S. Shamim, S. Cang, H. Yu, and Y. Li, "Management approaches for Industry 4.0: A human resource management perspective," in *2016 IEEE Congress on Evolutionary Computation, CEC 2016*, 2016.
- [6] H. Schroth, "Are you ready for gen Z in the workplace?," *Calif. Manage. Rev.*, vol. 61, no. 3, pp. 5–18, May 2019.
- [7] A. Alexandros and P. Dekoulou, *Cases on Corporate Social Responsibility and Contemporary Issues in Organizations*. Hershey, United States: IGI Global, 2019.
- [8] 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].
- [9] 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](https://www.huffpost.com/entry/give-them-what-they-want_b_8783712). [Accessed: 20-Jul-2019].
- [10] H. J. Anderson, J. E. Baur, J. A. Griffith, and M. R. Buckley, "What works for you may not work for (Gen)Me: Limitations of present leadership theories for the new generation," *Leadership Quarterly*, vol. 28, no. 1. Elsevier Inc., pp. 245–260, 01-Feb-2017.
- [11] E. K. Dizaho, R. Salleh, and A. Abdullah, "Achieving Work Life Balance Through Flexible Work Schedules and Arrangements.," *Glob. Bus. Manag. Res.*, vol. 9, no. Special Issue, pp. 455–465, 2017.
- [12] 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].
- [13] 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.
- [14] 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.
- [15] C. E. Ebeling, *An Introduction to Reliability and Maintainability Engineering*. Long Grove: Waveland Press, 2010.
- [16] M. Teimouri and A. K. Gupta, "On the Three-Parameter Weibull Distribution Shape Parameter Estimation," *J. Data Sci.*, vol. 11, pp. 403–414, 2013.
- [17] C.-D. Lai, D. N. Murthy, and M. Xie, *Weibull Distributions and Their Applications*. Springer, 2006.
- [18] 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.
- [19] 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.
- [20] 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.
- [21] R. N. Lussier and J. R. Hendon, *Human Resource Management: Functions, Applications, Skill Development*. SAGE Publications, 2016.