Predicting and enhancing employee daily performance for DEWA-HV design department

Maryam Abdulrahim Alshenaini
maa8474@rit.edu

Follow this and additional works at: https://scholarworks.rit.edu/theses

Recommended Citation

This Master's Project is brought to you for free and open access by RIT Scholar Works. It has been accepted for inclusion in Theses by an authorized administrator of RIT Scholar Works. For more information, please contact ritscholarworks@rit.edu.
RIT

Predicting and enhancing employee daily performance for DEWA-HV design department

By

Maryam Abdulrahim Alshenaini

A Graduate Paper/Capstone Submitted in Partial Fulfilment of the Requirements for the
Degree of Master of Science in Professional Studies: Data Analytics

Department of Graduate Programs & Research

Rochester Institute of Technology
RIT Dubai
May 16, 2020
RIT

Master of Science in Data Analytics

Graduate Paper/Capstone Approval

Student Name: Maryam Abdulrahim Alshenaini

Paper/Capstone Title: Predicting and improve employee daily performance for DEWA-HV design department

Graduate Paper/Capstone Committee:

<table>
<thead>
<tr>
<th>Name</th>
<th>Dr. Sanjay Modak</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chair of committee</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name</th>
<th>Dr. Ioannis Karamitsos</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Member of committee</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Acknowledgments

First, I would express my gratitude and appreciation to the supervisor of this study, Dr. Ioannis Karamitsos, for his unparalleled support and all guidance which he gave me.

Also, I would like to thank my manager Abdalla Al Hammadi, who was supportive, and his support in the process of collecting the data the study required.

I would also like to thank my parents very profoundly for loving and helping me during my Master's Journey research.

Finally, I would like to thank all those who helped me, gave me advice, and motivated me in my study.
Abstract

The most measurement tool used to measure daily employee performance is the daily productivity rate. It is an important indicator to understand employee daily performance and achievement. As a management level or as any successful manager, they are looking to achieve a higher percentage of daily productivity per employee to increase their department’s total outcome.

The primary purpose of this project is to implement the predicting and model and analysis techniques to improve and forecast the productivity of HV-design employees and reduce the lost time per day. Also, this project aims to identify the current situation of daily employee performance by using a productivity rate. Also, identify the leading causes that affect an individual's productivity, which leads to an increase in lost time. Determine the relationship between employee happiness and daily productivity rate and how happiness can affect the daily productivity rate. Moreover, using time series to build a predicting model to predict daily employee productivity for the next years. Finally, predict the best time of the year to let the employee goes to have different training courses. To achieve the project goals, I will depend on the dataset of the Dubai Electricity and Water Authority HV-design department from 2017 till November 2019. The dataset includes daily productivity and lost time per day.

Keywords: performance, daily productivity, employee, HV-design, predict, predicting model.
ACKNOWLEDGMENTS……………………………………………………………………………………………………………………………... II
ABSTRACT ....................................................................................................................................................................................... ERROR! BOOKMARK NOT DEFINED.
LIST OF FIGURES ..................................................................................................................................................................... ERROR! BOOKMARK NOT DEFINED.
LIST OF TABLES ........................................................................................................................................................................... ERROR! BOOKMARK NOT DEFINED.
CHAPTER 1 – INTRODUCTION ...................................................................................................................................................... 1
  1.1 INTRODUCTION .................................................................................................................................................................... ERROR! BOOKMARK NOT DEFINED.
  1.2 BACKGROUND ABOUT DEWA AND HV DESIGN APPROVAL .................................................................................................. ERROR! BOOKMARK NOT DEFINED.
  1.3 STATEMENT OF THE PROBLEM ............................................................................................................................................. 2
    1.3.1 THE IMPORTANCE OF THE STUDY ..................................................................................................................................... 3
    1.3.2 RESEARCH QUESTIONS ....................................................................................................................................................... 3
    1.3.3 RESEARCH HYPOTHESES .................................................................................................................................................... 3
  1.4 PROJECT GOALS ....................................................................................................................................................................... 3
  1.5 AIMS AND OBJECTIVES ............................................................................................................................................................ 4
  1.6 RESEARCH METHODOLOGY .................................................................................................................................................... 4
  1.7 LIMITATIONS OF THE STUDY .................................................................................................................................................... 6
CHAPTER 2 – LITERATURE REVIEW ............................................................................................................................................. 7
  2.1 INTRODUCTION ....................................................................................................................................................................... 7
    2.1.1 PRODUCTIVITY METRICS ................................................................................................................................................... 7
  2.2 IMPORTANT OF THE EMPLOYEE’S PRODUCTIVITY .................................................................................................................... 7
  2.3 THE RELATION BETWEEN THE EMPLOYEE PRODUCTIVITY AND EMPLOYEE HAPPINESS AND WORK ENGAGEMENT ................................................................................................................................................. 8
  2.4 HOW TO IMPROVE THE PRODUCTIVITY ................................................................................................................................... 8
  2.5 ADVANTAGES OF HIGH EMPLOYEE PRODUCTIVITY ................................................................................................................ 9
CHAPTER 3- PROJECT DESCRIPTION ............................................................................................................................................. 10
  3.1 DATA UNDERSTANDING AND PREPARATION ........................................................................................................................... 10
    3.1.1 DATA UNDERSTANDING AND EXPLORATION ....................................................................................................................... 10
    3.1.1.1 Total productivity and lost time per year ............................................................................................................................ 11
    3.1.1.2 The productivity per day for the last three years ................................................................................................................ 11
    3.1.1.3 Lost time per day for the last three years .......................................................................................................................... 12
    3.1.1.4 The productivity and lost time percentage of 2019 .............................................................................................................. 12
    3.1.2 DATA CLEANING ................................................................................................................................................................. 13
CHAPTER 4- DATA ANALYSIS ........................................................................................................................................................... 14
  4.1 FORECASTING MODEL ............................................................................................................................................................... 14
  4.2 FORECASTING MODEL TEST ....................................................................................................................................................... 16
    4.2.1 FIRST FORECASTING MODEL TEST (EMPLOYEE #1) ........................................................................................................... 16
List of Figures

Figure 1: DEWA’S organizational chart for HV design approval ................................................................. 2
Figure. 2: CRISP-DM Process ....................................................................................................................... 5
Figure. 3: The percentage of Productivity and lost time for last three years .............................................. 11
Figure 4: The productivity per day for the last three years ........................................................................... 11
Figure 5: Lost time per day for the last three years ...................................................................................... 12
Figure. 6: Daily performance of 2019 ........................................................................................................ 12
Figure. 7: The daily productivity for last years (2017 to 2019) .................................................................. 14
Figure. 8: Forecast plot ................................................................................................................................. 15
Figure. 9: Forecast components ..................................................................................................................... 15
Figure. 10: The daily productivity for last years (2017 to 2019) (employee #1) ........................................ 16
Figure. 11: Forecast plot for employee #1 .................................................................................................... 17
Figure. 12: Forecast components for employee #1 ..................................................................................... 17
Figure. 13: The daily productivity for last years (2017 to 2019) (employee #2) .......................................... 18
Figure. 14: Forecast plot for employee #2 ................................................................................................... 19
Figure. 15: Forecast components for employee #2 .................................................................................... 19

List of Tables

Table. 1: Variable Description ...................................................................................................................... 10
Table 2: Missing values of each variable ..................................................................................................... 13
Table 3: The new data frame with forecast object ..................................................................................... 14
Table 4: The new data frame with forecast object (employee #1) ............................................................ 16
Table 5: The new data frame with forecast object (employee #2) .............................................................. 18
Chapter 1- Introduction

1.1 Introduction

Employee productivity known as workplace productivity is an employee’s or a group of employees' efficiencies in a specific time according to Harness Jill (2018). In most situations, an individual's productivity will compare to the overall performance of other workers performing similar work.

Each employee's daily productivity is one of the most relevant results that managers and DEWA management generally care about it. The management level set a target that the employees should achieve 90% to 95% of the productivity per day. However, nowadays, they are never obtaining this percentage, which leads to affect the quality of the working hours and affect the department delivery time.

The purpose of this study is to analyses the employee's performance and explore the element that could impact their time management. Also, knowing and identify the challenges that the employee face, which could affect them from achieving higher productivity. Moreover, to have an insight into different factors that could increase the employee's daily productivity at the same time, it could also lead to an increase in the happiness rate of the employees.

1.2 Background about DEWA and HV design approval

Dubai Electricity and Water Authority (DEWA) is a company operated by the Dubai government. They were setting their strategy in line with Dubai strategy and UAE energy strategy to improve their services and providing innovative services for the customers. Besides that, the DEWA management looks out on their employees and set a lot of strategies to improve the employee performance and job satisfaction, such as employment happiness program and employment training program.
As shown in figure 1, HV design approval is part of the DEWA distribution power division. The HV design approval department deals with the high voltage substation room all around Dubai for different kinds of projects and even different types of substation rooms. The HV design department calculates each employee’s performance based on the total projects released per day; it can be either an approved project or return to the customers for modification or the submission of required documents. The target performance rate of the department was 80% to 85% for each employee per day. Nowadays, the design approval department managers what to raise the productivity target to 90% to 95% performance per day of each employee in the department. That has a higher impact on customer happiness because the employee will work more and approved the project faster. Moreover, one of the values of DEWA is stakeholders’ happiness; customers and employees are part of the stakeholders. So, by increasing both the customer and employee happiness and satisfaction rate, it will impact overall on the total percentage of stakeholders' happiness.

1.3 Statement of the Problem

The productivity report of HV-Design department starts in 2014, as part of SAP system services. And each employee submits his/her daily productivity. Meanwhile, the managers set a target to achieve 90% to 95% of daily productivity for each employee. But they observe that the productivity rate is low and not even close to their target. So, they start to hire more employees from the third quarter of 2015 tell 2018. In order to achieve their objective. However, the issue is still not solving. This project aims to provide valuable information and clear understating about daily productivity.
1.3.1 The importance of the study

The study of employee's daily performance and productivity is of great importance because it affects the company, and it affects the fate of every individual that wants to develop in their career.

Furthermore, predicting and understanding the employee's daily productivity and knowing each employee's capability will help and support managers in developing improvement strategies for each employee and finding the best programs and training courses to improve their daily performance.

This study is the first study that uses and deals with the daily employee productivity for Dubai Electricity and Water Authority employees, especially HV design approval employees.

1.3.2 Research Questions

1- How was the performance for the last three years?
2- Can we predict next year's performance?

1.3.3 Research Hypotheses

Through this study, the researcher's hypotheses focus on increasing the daily productivity rate of each employee in the HV design approval department.

1.4 Project goals

One of the main strategic goals of Dubai and DEWA 2021 strategy is to provide a happy and positive work environment (DEWA (2016)). From this strategy goal, I will focus on this project on the impact of employee's happiness and satisfaction level on the productivity rate. Also, this study will investigate the challenges that the employee faces daily during the working hours. And visualize the current situation of productivity using visualization tools. Moreover, one of the project objectives is to develop a methodology of predicting a model to predict the next year's performance and productivity rate of each employee and predicting the best time to take training courses during the year, based on the analyzing the datasets. Finally, the project aims to reduce the lost time per day to increase productivity per day.
1.5 Aims and Objectives

This study aims at the following:

1. To determine employee productivity level in the HV dosing approval department, during the period 2017 to 2019.
2. Identify the average percentage of daily lost time in working hours.
3. Forecast the productivity rate for the next period based on the registered historical data from 2017 to 2019.

1.6 Research Methodology

The productivity dataset consists of the productivity rate and lost time of each employee in the DEWA HV-design department from 2017 to November 2019. I will explore the relationship between employee productivity lost time per day. And explore the challenges that employees were facing daily.

The data were obtained from the HV design approval Department, as raw basis data from written reports. Also, we took the approval from Approval design department manager for using the data in this study. The data was feed into the Python (open sours tool), to understand and know available data. We decided to deal with data of one employee in the HV department and build a prediction model based on this employee data. Then, we will apply this model to other employees.
Moreover, we decided to use and follow the CRISP-DM process, which includes five stages as depicted in the following Figure 2:

1. Business understanding
2. Data understanding
3. Data preparation
4. Modeling
5. Evaluation and Deployment.

First, I collected the productivity dataset from the DEWA employee portal on productivity records. Using open sources tool Python, I cleaned the data and prepared the dataset, to find results. Also, I removed unnecessary attributes from the dataset. I used quantitative theory to get the correlation between the variables in the productivity dataset. Afterward, I visualized all the results I got. Then, I split the data into two datasets: training data and testing data, and building predicting models for next year’s predictivity performance based on one employee’s data. And evaluate the results.
1.7 Limitations of the Study

This study includes the Daily productivity report for the HV design department. This dataset as a lot of missing value and repeating columns. So, it was challenging for me to deal with the dataset to find the results. And about the timeframe, this study will cover the record productivity percentage data from 2017 to 2019, for all employees. Moreover, we planned to get the employee happiness survey results of the HV design approval department. Still, the responsible department was not allowed us to get it and study it.
Chapter 2 – Literature Review

2.1 Introduction

There is quite a little research that has been done about employee daily productivity, either in a local organization or globally. Therefore, these studies will be covered and reviewed to identify the results of these studies. To support this research paper, with similar beforehand studies results.

2.1.1 Productivity metrics

Productivity is a crucial part of the employee’s performance in any organization. It refers to the amount of work that an employee completes in a particular time frame, such as fixed working hours. Managers often measure productivity as the number of projects reviewed by employees per day (RAJINI, K., and M. PANIMALAR;).

2.2 Important of the employee’s productivity

According to Stevens (2017), one of the irritating things for managers and supervisors is to have employees that never achieve the department goal for daily productivity. So, it is a clarification of the causes behind bad results is essential for managers. Because these issues of not achieving the productivity goal will still be an issue if the managers do not make a diagnosis of the problem to have low productivity, accordingly to that, the manager and supervisors must set a plan to understand this issue and apply the rationale for improvement this issue. Moreover, it will give a chance for the managers and supervisors to knowing their employees’ performance per day.
2.3 The relation between the employee productivity and employee Happiness and work engagement

Zelenski, Murphy & Jenkins (2008) mention that there is still uncertainty about whether the happy workers are actually more productive. Also, they mention that by knowing the employee's happiness level, management can set a clear strategy to improve employee happiness and productivity.

Stevens (2017) mentions one of the reasons that affect employee happiness level is that there is missing clear work goals and objectives, as is shown in the survey that 67% of employees expressed unhappiness demonstrated to them because of that reason. Also, it mentions it is important to know if there is any other impact of employee happiness in the workplace on employee productivity.

Singh.SH, David.Raj & Mikkilieni.Sit (2018) indicate that enhancing employee happiness is an essential means of improving an organization’s performance efficiency. Also, refer to Singh.SH, David.Raj & Mikkilieni.Sit (2018) indicate that employee happiness is an essential outcome for both employees and organizations that want to achieve or access their human capital to its full capacity. Moreover, they mention that feeling positive most often has an effect on happiness partly overlaps with defining work engagement. But the significant distinction between happiness and work engagement is that happiness is connected with the context-free and general psychiatric state while work engagement is precise and career-related.

However, Ledford Jr & Gerald E (1999) stated in their research, the hypothesis of happiness – productivity required a person test. Because he believes that, perhaps employee unhappy, but he/she tends to be more productive. So, it is crucial to understand the behavior of the employee and when he/she be more productive.

2.4 How to improve the productivity

Stevens (2017) mentions that the employee lacks the skills required. The employees have to be responsible for their respective duties, but they must be trained to ensure they will be comfortable to deal with job tasks and be more familiar with the work. Accordingly, one of the steps to improve productivity with a set clear training plan for each employee.
Tulgan & Bruce (2017) researcher provides four tasks for managers and doing one task at a time: the first task is monitoring the employee while they are doing the task multiple times. Then divide each main task into steps and break each step into a series of actions to make it easier for the employee and finish the task faster. The second task is to discover how long the job will take to be done and develop a time budget for each task to complete the main work and identify the specific time required for each task. The third task is to do the micro-gap analysis to indicate the difference between the time budget set for the task and the employee's actual time need to finish the task.

Furthermore, these micro-gaps will have potential opportunities to improve and speed up work because it will give a clear understanding of the required time to spend to finish each task and the last task is to choose one task at a time and take it slowly to understand it more. Then fill the micro-gaps step by step. When it related to helping an employee, you have to speed up, here are the questions you have to look in it:

- Is the employee following best practices or not?
- Is the employee doing unneeded tasks while doing their work?
- Is the employee setting useless steps in any required tasks?
- Is the employee confronting any repeating difficulties that have not paid attention to it and not taking it seriously?

2.5 Advantages of high employee productivity

Improving employee productivity can lead to many advantages for either individual or companies. The first advantage is achieving the company's goals or department; if the employee is more productive, that means there is an improvement in the quality and quantity of their work. That leads to having a higher output from the employee daily and achieving the department goals as a result (Collier, Ellie (2018)).

According to Harness, Jill (2018), an employee with highly productivity rate will provide accelerated and higher-quality customer service than other employees. It enhances the relationship between the customers and the companies. Outstanding customer service can positively affect customer loyalty, and that can lead to an increase in the company's profit.
Chapter 3- Project description

3.1 Data understanding and preparation

3.1.1 Data Understanding and Exploration

In this section, we will do an exploration of the collected dataset to understand it. The data for this project was collected through the DEWA eForm data for the HV design approval department as a separate excel sheet file for each employee's productivity. The collected data was for the last three years, including seven employee's daily performance. This data was divided into two datasets. The first dataset of 760 records for one employee was used for training and testing, while the other dataset for the rest of the employees was used for validating the model. The initial dataset is a data frame of one employee having 758 objects and 11 variables.

<table>
<thead>
<tr>
<th># of variable</th>
<th>Variable</th>
<th>Description</th>
<th>Data type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Day</td>
<td>The working day: Sunday to Thursday</td>
<td>Text</td>
</tr>
<tr>
<td>2</td>
<td>Date</td>
<td>Date, DD-MM-YYYY</td>
<td>Date</td>
</tr>
<tr>
<td>3</td>
<td>Year</td>
<td>Year of summited report in the system. It is between 2017 till November 2019</td>
<td>Continuous number</td>
</tr>
<tr>
<td>4</td>
<td>Submitter Name</td>
<td>The employee names.</td>
<td>Text</td>
</tr>
<tr>
<td>5</td>
<td>Total Hours Available</td>
<td>Working hours per day</td>
<td>Number</td>
</tr>
<tr>
<td>6</td>
<td>Total Allowed Travel Time</td>
<td>The available travel or traffic hours per day that not used.</td>
<td>Number</td>
</tr>
<tr>
<td>7</td>
<td>Total Earned Hours</td>
<td>The total of actual working hours per day</td>
<td>Number</td>
</tr>
<tr>
<td>8</td>
<td>Total Lost Time</td>
<td>The known causes for lost time, such as preparing to start work, breakfast and prayer.</td>
<td>Number</td>
</tr>
<tr>
<td>9</td>
<td>Total Non-Identified Lost Time</td>
<td>The unknow reasons for lost time.</td>
<td>Number</td>
</tr>
<tr>
<td>10</td>
<td>Productivity</td>
<td>The percentage of total productivity per day.</td>
<td>Number</td>
</tr>
<tr>
<td>11</td>
<td>Lost Time</td>
<td>The percentage of total lost time day.</td>
<td>Number</td>
</tr>
</tbody>
</table>

Table 1: Variable Description

Table.1 includes all attributes of the data with description and data type of each attribute. As shown in the table, there are eleven attributes in the dataset: day, date, year, submitter name, total hours available, Total Allowed Travel Time, Total Earned Hours, Total Lost Time, Total NonIdentified Lost Time, Productivity and Lost Time.
3.1.1.1 Total Productivity and lost time per year

We will visualize the average productivity percentage per day for one employee for the last three years. As per the figure 3, all the three years of almost had the same productivity rate. While the lost time rate slightly decreases in 2018 compared to 2017 and 2019. And it is somewhat increased in 2019 compared to the past two years.

![The percentage of Productivity and lost time for last three years](image)

Figure 2: The percentage of Productivity and lost time for last three years

3.1.1.2 The productivity per day for the last three years

We notice from Figure 4 that the employee submitted the productivity report almost for the full year of 2017. However, there are a lot of missing productivity records in both 2018 and 2019. Specifically, the year of 2019, has the lowest record compare to the record data of both 2017 and 2018. From that, we can understand that the performance level of the employee is still the same for the last three years. And it could the employee does not submit his / her productivity in DEWA eForm one of the reasons why it is not recorded it.

![Productivity for last three years](image)

Figure 4: The productivity per day for the last three years
3.1.1.3 Lost time per day for the last three years

From Figure 5, we noted that the employee submitted the daily report almost for the full year of 2017. The productivity chart outlines a lot of missing productivity records in both 2018 and 2019.

Figure 5: Lost time per day for the last three years

3.1.1.4 The productivity and lost time percentage of 2019

Figure 6 shows that the distribution of daily performance for the year 2019 is divided into productivity and total lost time per seven hours, regular working hours. We notice that some days have a total working hour is five, while the rest seven hours. We foresee that these days are the working days in Ramadan because the working hours will be five hours. We even note that there are a few days where the lost time percentage is higher than the productivity percentage.

Figure 6: Daily performance of 2019
3.1.2 Data Cleaning

The primary dataset comprises 758 records of the daily performance of one of the employees for the last three years (2017 to 2019). Initially, we found that there are a lot of missing values for each variable in the data, as indicated in the Table 2. Table 2: Missing values of each variable

<table>
<thead>
<tr>
<th>Variables</th>
<th>Number of missing values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day</td>
<td>0</td>
</tr>
<tr>
<td>Date</td>
<td>0</td>
</tr>
<tr>
<td>Year</td>
<td>0</td>
</tr>
<tr>
<td>Submitter Name</td>
<td>0</td>
</tr>
<tr>
<td>Total Hours Available</td>
<td>282</td>
</tr>
<tr>
<td>Total Allowed Travel Time</td>
<td>282</td>
</tr>
<tr>
<td>Total Earned Hours</td>
<td>282</td>
</tr>
<tr>
<td>Total Lost Time</td>
<td>282</td>
</tr>
<tr>
<td>Total Non-Identified Lost Time</td>
<td>282</td>
</tr>
<tr>
<td>Productivity</td>
<td>282</td>
</tr>
<tr>
<td>Lost Time</td>
<td>282</td>
</tr>
</tbody>
</table>

According to that, we were deciding that we will remove all rows with missing values. Based on that, the total record becomes 476. Moreover, we choose to drop unnecessary columns that we will not use it in the analyses from dataset such as submitter name, day, total allowed travel time. After the data preparation, the data frame consists of 476 records and eight variables.
Chapter 4- Data Analysis

4.1 Forecasting Model

For this study, we used a time-series method to discover the pattern of the historical data, and then we estimated the pattern record of the future data. In this section we used the training data that we prepared and clean it, as presented in the previous chapter 3.

Predictions are made on a data frame that includes column DS. Which including the dates for which a prediction is to be made. It would also have the historical periods by default, and the model has been fitted. The predicted model would add an expected value to each row, which it names YHAT. YHAT is a future predicted value that assigned by the predict method for each row and it will provide an in-sample fit. The rang of this value defined by yhat_lower and yhat_upper. The forecast here is a new data frame that includes a YHAT variable with the forecast and component columns and uncertainty intervals.

Table 3: The new data frame with forecast object

<table>
<thead>
<tr>
<th>DS</th>
<th>YHAT</th>
<th>YHAT_LOWER</th>
<th>YHAT_UPPER</th>
</tr>
</thead>
<tbody>
<tr>
<td>2020-12-04</td>
<td>85.032881</td>
<td>74.072983</td>
<td>96.232650</td>
</tr>
<tr>
<td>2020-12-05</td>
<td>84.961913</td>
<td>73.888605</td>
<td>96.624759</td>
</tr>
<tr>
<td>2020-12-06</td>
<td>84.041374</td>
<td>72.636805</td>
<td>96.470291</td>
</tr>
<tr>
<td>2020-12-07</td>
<td>83.415512</td>
<td>72.695483</td>
<td>94.547567</td>
</tr>
<tr>
<td>2020-12-08</td>
<td>83.659097</td>
<td>71.150302</td>
<td>95.293806</td>
</tr>
</tbody>
</table>
Table.3 shows the last five rows of the new data frame that include the forecast object as YHAT, component columns and uncertainty intervals.

Figure.8 shows a new data frame that includes a YHAT variable with the forecast and component columns and uncertainty intervals. The orange line is the divider between the record data of the last three years. And after the orange line, it is the predicting record for the year of 2020. It is shown that the performance level of the employee will have almost the same performance as in 2019.

![Forecast Plot](image)

**Figure 8: Forecast plot**

Figure.9 shows the forecast component plot, which including the trend, yearly seasonality and weekly seasonality of the time series by default.

![Forecast Components](image)

**Figure 9: Forecast components**
4.2 Forecasting Model test

In this section we used two different datasets for two employee performance record for the last three years 2017-2019. We prepared both datasets as we did for the training data.

4.2.1 First forecasting model test (employee #1)

After remove all rows with missing values, the total record becomes 213. Also, we choose to drop unnecessary columns. We end with a data frame of 213 records and eight variables.

It is clear from figure 10 that this employee maintained the performance rate between 80% to 85% for the year 2017 and 2018. And we note that there are no records for 2018 performance for this employee. It is difficult to know why that happens in 2018, but one of the reasons it could be is that the employee forgotten or not feel the importance of submitting his / her daily performance report.

![Productivity for the last three years (2017 - 2019) Employee#1](image)

Figure 10: The daily productivity for last years (2017 to 2019) (employee #1)

Table 4: The new data frame with forecast object (employee #1)

<table>
<thead>
<tr>
<th>Date</th>
<th>YHAT</th>
<th>YHAT_LOWER</th>
<th>YHAT_UPPER</th>
</tr>
</thead>
<tbody>
<tr>
<td>2020-12-01</td>
<td>85.266361</td>
<td>83.137233</td>
<td>87.449899</td>
</tr>
<tr>
<td>2020-12-02</td>
<td>84.788687</td>
<td>82.698071</td>
<td>87.027047</td>
</tr>
<tr>
<td>2020-12-03</td>
<td>84.595061</td>
<td>82.617924</td>
<td>86.302744</td>
</tr>
<tr>
<td>2020-12-04</td>
<td>84.206709</td>
<td>82.210961</td>
<td>86.367364</td>
</tr>
<tr>
<td>2020-12-05</td>
<td>84.221278</td>
<td>82.198725</td>
<td>86.367364</td>
</tr>
</tbody>
</table>

Table 4 shows the last five rows of the new data frame.
Figure 11 shows a new data frame for employee #1. After the orange line, it is the predicting record for the year of 2020. On the other hand, it seems that the model was predicting the history data of 2018 since the employee did not have any document for this year. Moreover, it is noted that the performance of this employee will be improved in 2020.

Figure 11: Forecast plot for employee #1

Figure 12 shows the forecast component plot for employee #1, which including the trend, yearly seasonality and weekly seasonality of the time series by default.
4.2.2 Second forecasting model test (employee #2)

After remove all rows with missing values and drop unnecessary columns. We end with a data frame of 382 records and eight variables.

From the figure 13; we understand that the employee maintained the performance rate between 80% for the year 2019.

![Productivity for the last three years (2017 - 2019) Employe#2](image)

Table 5: The new data frame with forecast object (employee #2)

<table>
<thead>
<tr>
<th>DS</th>
<th>YHAT</th>
<th>YHAT_LOWER</th>
<th>YHAT_UPPER</th>
</tr>
</thead>
<tbody>
<tr>
<td>2020-12-06</td>
<td>71.713525</td>
<td>66.714665</td>
<td>77.032411</td>
</tr>
<tr>
<td>2020-12-07</td>
<td>70.651204</td>
<td>65.585249</td>
<td>76.311786</td>
</tr>
<tr>
<td>2020-12-08</td>
<td>70.189924</td>
<td>65.035073</td>
<td>75.490522</td>
</tr>
<tr>
<td>2020-12-09</td>
<td>69.928355</td>
<td>64.835335</td>
<td>74.925803</td>
</tr>
<tr>
<td>2020-12-010</td>
<td>69.725630</td>
<td>64.374859</td>
<td>75.136950</td>
</tr>
</tbody>
</table>

Table.5 shows the last five rows of the new data frame.

Figure 14 shows a new data frame of employee #2. After the orange line, it is the predicting record for the year of 2020. As is noticeable, employee performance will decrease significantly in 2020.
Figure 14: Forecast plot for employee #2

Figure 15 shows the forecast component plot, which including the trend, yearly seasonality and weekly seasonality of the time series by default.

Figure 15: Forecast components for employee #2
Chapter 5- Conclusion

5.1 Conclusion

This study focused on the possibilities of building a model to forecast next year’s daily performance of the employee. During working with the performance dataset, many attributes have been tested, and some of them are found useful on the performance prediction. The productivity rate was the strongest attribute. Because of that, we used the productivity attribute for building the forecasting model. The total record of the training data was 758 with 11 attributes for the last three years 2017-2019. After removing all missing values and unnecessary columns, they became 476 with eight attributes. And with testing the model for two other employees, we realized that even in the year 2020, the employee would not be achieving the department target of 90%-95% of daily productivity.

Notwithstanding, this model can be used for the human resources departments to enhanced and develop a strategy to improved employee performance per day. Also, it can be used for predicting new employee’s performance, which leads to putting an action plan to avoid any risk related to low performance. Make it more advanced by adding new variables such as public holidays, training days and manager feedback.
5.2 Recommendations

For this study, I recommend collecting more related data from the employee happiness department and another department and find the correlation between the happiness level of employee and job satisfaction on the daily performance of the employee. It can also use the database of a different employee from different departments to get a correct performance rate for each employee. Also, we recommend developing a model that includes the rules of the human resources department to predict employees’ performance based on the HR laws.
References


7- Stevens, Drew. "increase employee productivty without additional time and resources." 2017.


9- VK, Anirudh. *How The CRISP-DM Method can help manage your next data science project*. n.d.