EEET2485 – RESEARCH METHODS FOR ENGINEERS

GROUP PROJECT: DATA ANALYSIS

Group number: 1

<u>Lecturer</u>: Dr. Son Dao Vu Truong

<u>Students:</u> Nguyen Thanh Luan - s3757937

To Vu Phuc – s3758272

Nguyen Phuong Uyen – s3751882

Nguyen Le - s3777242

Submission date: 29/05/2021

We declare that in submitting all work for this assessment we have read, understood, and agree to the content and expectations of the Assessment Declaration.

Contents

1. Abstract:

The garment industry is among the most labor-required occupations for its nature of little reliability on machinery and automation. For the significant demands on human work, final products' quality from garment manufacture is of dependency on target of productivity. This research takes on investigation and employment of the impacts on the garment workers and their productivity. With the given dataset, methods of posing questions on relationships among the factors and data analysis are prompted to find and prove the hypotheses. For supporting the measures, statistical analysis and data mining techniques are applied including Feature Selections and Apriori Algorithm to find the relationships as well as form up prediction and development trends. Overall, some rules are derived that can set further improvement to the actual productivity. After performance data analysis, some key points are drawn out:

- Actual productivity shows a strong influence from target productivity, which is initially set for the workers to accomplish per day.
- Incentive though has a low matching relationship, still shares a correlation as a driving factor proportional to the employees' performance.
- Idle men and idle time are also factors of impact on the productivity of the garment factory

From which results, the model is successfully generated to make a prediction and show prominent performance for future improvement of the garment workers' productivity.

2. Introduction:

Bangladesh is the 2nd largest ready-made garments (RMG) exporter in the garment industry [1]. In 2020, its export value of garments was $38.73 billion against the $50 billion target set by the government and authorities. In terms of production efficiency, its performance is lower than the mark which poses a critical issue amongst heavyweight competitors such as China, followed by German, Viet Nam & India. Under the context of a Bangladesh garment factory, the dataset is shown with several attributes of influence on productivity, whose relationships would be exploited through a methodology to answer hypothesis questions, with Python being the main language for data analysis. With the given dataset, a set of methods and algorithms are conducted for data analysis and relate the findings to the posed questions. Evaluation of data would be compared and deeply analyzed for results confirmation.

3. Methodology:

We follow the traditional data analysis steps, as shown in Fig 1, to conduct our study.

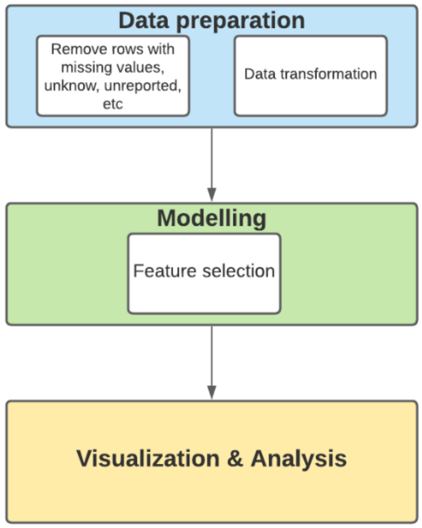


Figure 1. Methodology procedures

Data preparation

The dataset of study is the Productivity Prediction of Garment Employee (PPGE) from UCI Machine Learning Repository [2], originating from the industrial engineering department of the manufacturing unit. The dataset provides the quarterly data regarding garment company productivity from Jan 2015 to Mar 2015. Specifically, the PPGE dataset consists of 15 attributes and 1197 specific records of each. The focus of the dataset is the actual-recorded productivity of the garment manufacturing workers, and the target productivity stems from the reference of workers' accomplishment and other parameters that could serve as reliability elements of influence to the targeted values. For a final notice, most attributes in PPGE Dataset are numeric types, whereas 20% of columns are categorical types.

Dataset attributes clarification: Garments_worker_productivity.csv

Attributes	Entries	Description	Example attribute
Date	1197 entries	Date in MM-DD-YYYY	1/1/2015
Quarter	1197 entries	A portion of a month.	Quarter 1
Department	1197 entries	Associated department.	Sewing, finishing
Day	1197 entries	Day of the week	Thursday
Team	1197 entries	Associated team number.	Team 8, team 6
Target productivity	1197 entries	Targeted percentage of productivity set by the authority for each team per day.	0.8
SMV	1197 entries	Standard Minute Value – allocated time for a task	26.16

WIP	691 entries	Work in progress – number of unfinished items	1108
Overtime	1197 entries	Amount of overtime per team by minutes	7080
Incentive	1197 entries	Amount of financial incentive (in BDT) motivating a particular course of action.	98
Idle time	1197 entries	Amount of time when production is interrupted.	0
Idle men	1197 entries	Number of idle workers due to interrupted production	0
Number of style change	1197 entries	Number of changes in the style of a product	0
Number of workers	1197 entries	Number of workers per team	59
Actual productivity	1197 entries	The actual percentage of productivity accomplished by workers, ranging from 0-1	0.940725424

Table 1: Main group attributes

Data cleaning:

For getting the dataset into preparation before going through analysis, the content of the dataset is first clarified whether it is equivalent to the source document or against the information provided by the author in Kaggle and his website until the data of both sides is well-matched with sufficient evidence. Some information such as total entries (row), attributes (columns), and the type of each column are thoroughly checked. Common errors such as missing values, out-of-range values, typos, outliers, and bad lines from the given dataset are fixed to match each value of the dataset's contents. Specifically, this PPGE dataset only requires adjustment and handle in unreported values or missing values (NaN values). For instance, for solving the missing rate of up to 42% of the WIP attribute, the first solution coming up was to delete all the rows having missing values to avoid bias as much as possible. However, this is undesirable since the data frame was reduced to half compared with the original, so the problem was solved by applying the supervised learning algorithm k-nearest-neighbor (kNN) to fill the missing data.

Feature encoding:

Since there are three categorical features named "department", "team_no," "quarter," these attributes must be encoded before putting into the Python machine learning model. One-hot encoding techniques are applied to create binary columns, indicating each presence of original data, from which a total of 20 new features were obtained to feed into the model.

Research objectives and questions:

From the heatmap built-up of all attributes, hypothesis questions are posed as metrics to the analytical process.

- **Question 1**: What is the relationship between the actual productivity and the targeted one in both the sewing and finishing departments from 1 January 2015 to 11 March 2015?

- **Question 2:** Do the incentive and the actual productivity share a correlation? Is the incentive proportional to the overtime or the number of workers (potentially show workers' interest in their rewards)?
- **Question 3:** How do idle time and idle men influence actual productivity? What are possible factors for the cause of the high number of these two attributes?
- **Question 4:** Would a possibility of higher style change affect the actual productivity as well as the eventual profit?
- **Question 5:** How to make **prediction** and **improvement** if the pattern of the actual productivity is not evenly distributed for the declining trend of monthly productivity?

Data exploration and Modeling:

This process comprises some statistics and uses several kinds of graphs to show the basic characteristics of the dataset. We applied some machine learning algorithms such as Apriori and Decision tree to find out the underlying patterns and relationship attributes. Our results consist of a collection of rules derived from the machine learning model and summarized data from statistics.

Success metrics: Since we have a collection of rules related to the productivity of garment employees, our success metrics are practicality and specificity. If our rules are clear, useful, and easy to implement in real life, our research meets objectives.

4. Evaluation and Findings:

Statistic Results:

PPGE Dataset overview:

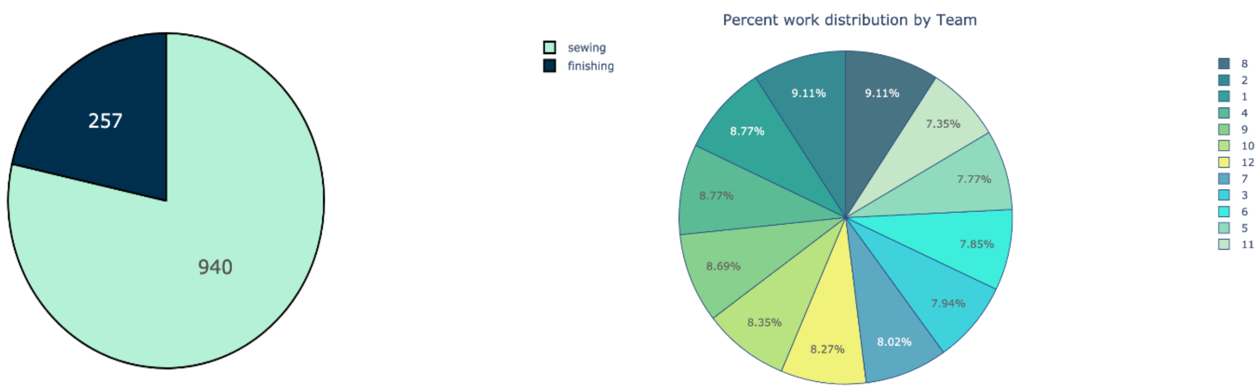


Figure 2. Overview of teams and department work distribution

In total 1197 entries of departments and garment teams, department of the on-progress takes up around 80% of total departments, much higher in terms of value compared to the finishing ones. Meanwhile each team out of 12 shares a moderately same workload, around 8-9% per team. In terms of attributes correlation shown in the heatmap of Figure 3, the co-dependence of some attributes presented to be the highest is man versus time (number of workers - overtime) sharing 0.73 and time allocation for each work (Standard Minute Value – overtime), which stays true to the working time of each worker extending accordingly to the estimated time spending on a particular task. The actual productivity, contrary to the initial prediction, is not related to the number of workers in a day, for both showing an extremely low match. Evaluating the relations between actual productivity and number of styles change, idle men, idle time, the recorded values respectively are -0.21, -0.18, -0.081, showing all as negative, which means these parameters are inversely proportional to the actual performance of garment workers. As in the factual scenario, the productivity is of major dependence on the workers' labor, therefore, there should be as little idle men and idle time as possible. Similarly, in the case of style changes, more accomplishments are produced if the process requires the least changes in garment designs, as the shifting phases would take up the efficient time of working.

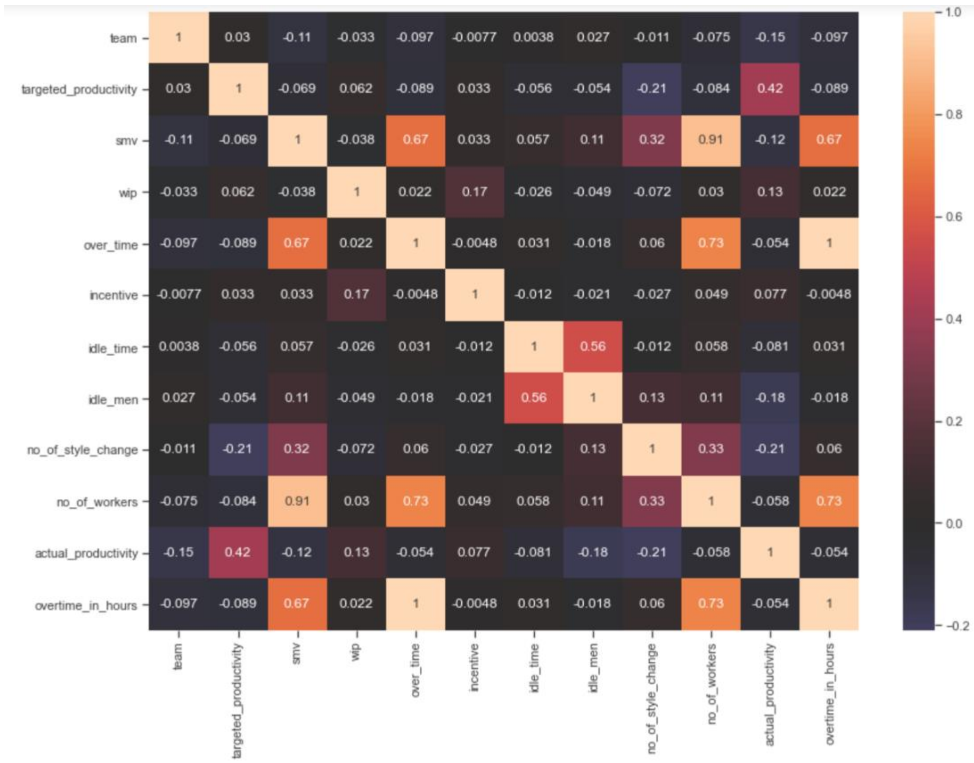


Figure 2. Heatmap of all attributes in the PPGE dataset

Separately considering the overall actual productivity of the garment employees on weekdays, which is shown in Appendix 3, Saturday appears as the most productive with average productivity of 0.75 because in Bangladesh, Friday is the weekly holiday. Most of the non-governmental organizations have Friday as their only holiday per week, so after a break productivity should be boosted. Similarly, during the first month of the year - January, would their average working performance show a more positive perspective compared to the later months.

Question 1: What is the relationship between the actual productivity and the targeted one in both the sewing and finishing departments from 1 January 2015 to 11 March 2015?

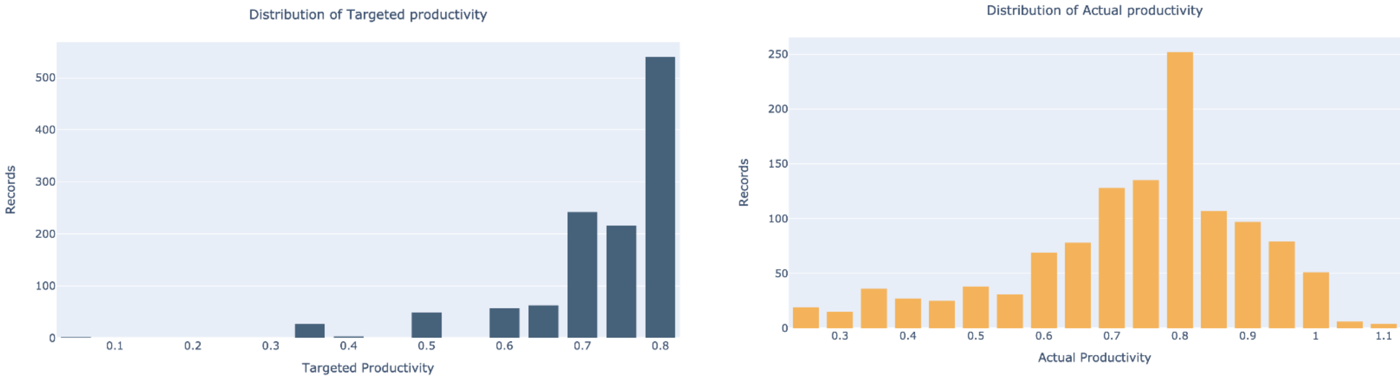


Figure 3. Targeted productivity (left) and actual productivity by total records

The data recorded was from the given timeframe of the dataset – from 1st January to 11th March 2015 of expected productivity and actual productivity. The Authority set a daily targeted value for each team with a specific number of tasks to deliver at the end of each day. As in Figure 4, most were of around 70% to 80% productivity to achieve the daily target, only a minor amount of work was listed at a lower efficiency than 50%. On contrary to the initial criteria, the employees' productivity varied with main distribution among 65% to 90%, shown in the similar number of records – around 800 records in the mentioned ranges. Overall showed a more evenly distributed work performance in the actual employee productivity per day. Noticeably mentioned, the higher productivity of both cases was 80%, but in the real working scenario, this number was only achieved with half of the recorded days expected from the Authority. For the actual productivity, the garment workers showed a trend of their working efficiency to be dependent on the number of tasks as well as the working team, as differently assigned in each day workload.

To summarize the daily productivity of the garment workers, Appendix 1 and 2 showed that around 27% of the total records of actual productivity observed a lower working performance compared to the targeted one designated from the Authority.

Question 2: Do the incentive and the actual productivity share correlation? Is the incentive proportional to the overtime or the number of workers for these parameters potentially showing workers' interest in their rewards?

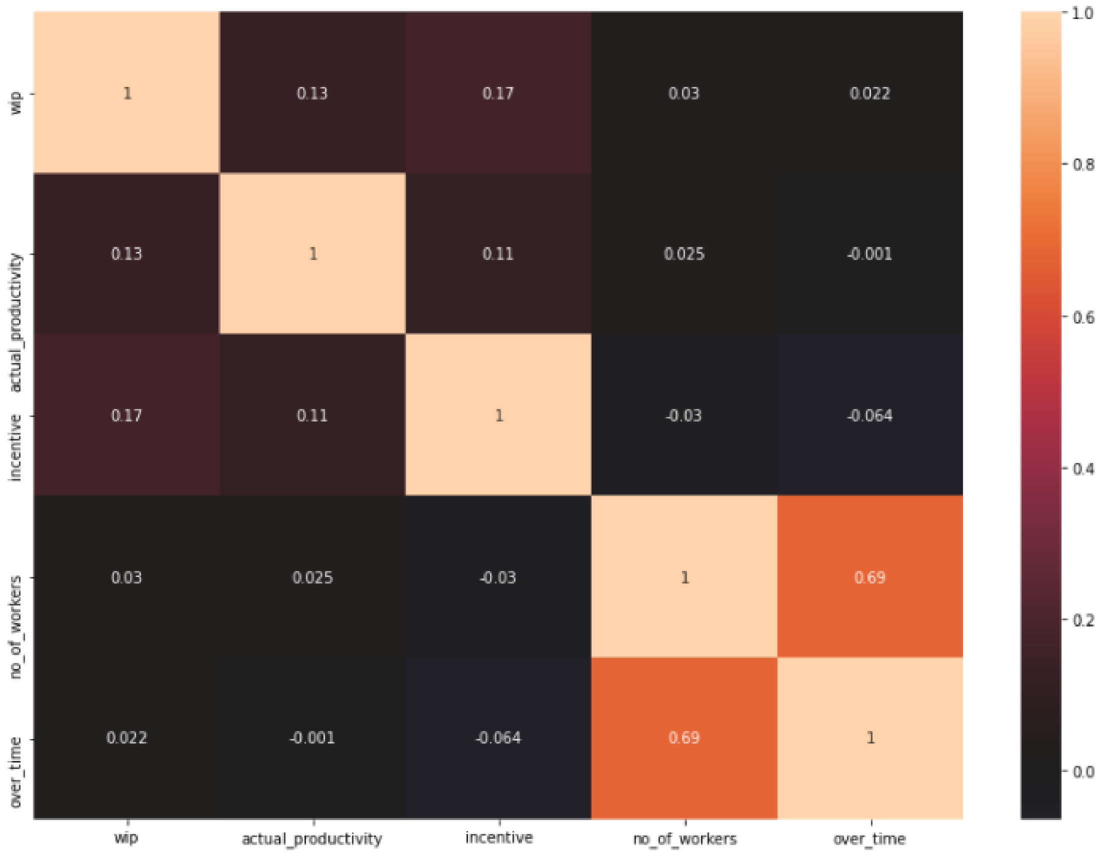


Figure 4: Correlation between incentive and WIP, no_of_workers, overtime and actual productivity of sewing department.

As shown in the matrix in Appendix 4 showing relationships between the incentive to actual productivity and WIP by departments, the finishing department does not have work in progress as well as incentives in their works. The third column of figure 5 also shows a correlation between incentive and actual productivity, at around 0.11. The incentive is higher when the actual productivity raises in the range of 20% to 100% on average. The incentive value is also correlated with the WIP value in the record. These recorded numbers reflect that the incentive encourages the garment teams to deliver more products even the standard working time is insufficient to finish within the same day. This helps the working departments to boost their productivity by completing higher portions of the work for the next day.

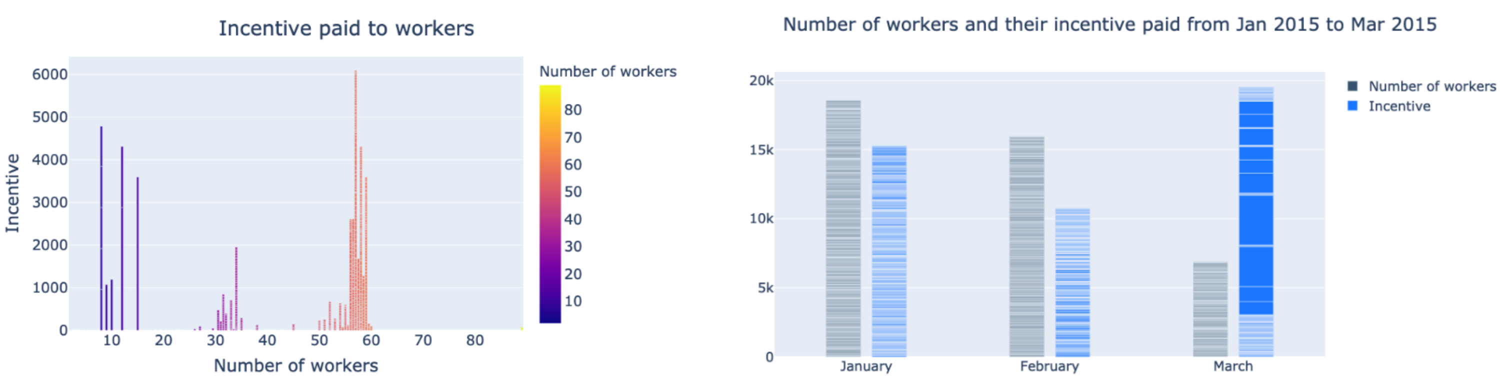


Figure 5: Incentive paid to workers in total (left) and in Jan, Feb & Mar

From figure 6, the incentive is paid in greater amount for the team with a higher number of workers. However, from the illustration of incentive paid for workers in separate three months, even there are fewer workers in March, the amount of incentive paid for workers observes a distinctly higher value. To explain this case, the correlation between WIP, actual productivity, incentive and overtime is prompted, shown in Appendix 5. These heatmaps show the correlation between attributes in each month that does not account for finishing departments because they do not have any incentive and WIP mentioned as in Appendix 4. In January, the incentive has the highest correlation with actual productivity in three months despite having no correlation with overtime. While in February, there is a slightly decreased correlation between incentive and actual productivity but an increase in correlation between incentive and overtime. Contrastingly in March, even though there is no correlation between incentive and actual productivity and even a negative correlation between incentive and overtime, garment employees in this month still achieve a higher amount of incentive with their poorer performance in productivity. From these numbers, we can conclude that there is no correlation between incentive and worker's interest in work as in the hypothesis question.

"An empirical study [3] suggested; reward system filled in as an impetus to make workers motivated and expand efficiency, they should be rewarded for their challenging work and punctuality. For example, award for attendance. Work coordination is likewise vital to expand profitability and meet generation target. Sound worker conduct and proficiency reward additionally meet the objective and build efficiency."

Study showed in [4] that ineffective management is the most significant factors that have strong impact a on the less productivity followed by outdated system, inadequate monetary and non-monetary rewards, Unsafe working conditions and the insufficient and ineffective coworkers."

Question 3: How do idle time and idle men influence actual productivity? What are possible factors for the cause of the high number of these two attributes?

Figure 6: Idle men and idle time vs. actual productivity

The number of idle men in the total report is seen mostly around 20 to 35 people, while the idle time does not concentrate on any area. With such recorded values, the actual productivity delivered by garment workers in these recorded entries, expectedly, is low for half of the record. The highest value of actual productivity in the idle men-concentrated region is 60%, achieved with around 25 idle workers. For the idle time above 100, the employees' performance could not reach 35% of working efficiency. When workers are habituated to unnecessary talking, this leads to more garment alterations and rejections. Because they pay less attention to their work and do not check the sewn garment properly before passing it to the next process operator. He/she is also negligent in checking the previous process before passing the garment to the next operator. Finally, when the garment reaches to inline quality checker many alterations are found & the garment is sent back to the respective operators for corrections. Frequently more than one operation is unfolded to make the correction which affects the productivity of that sewing line. Sometimes expert & mindful workers get disturbed by the noise created in the line.

Question 4: Would a possibility of higher style change affect the actual productivity as well as the eventual profit?

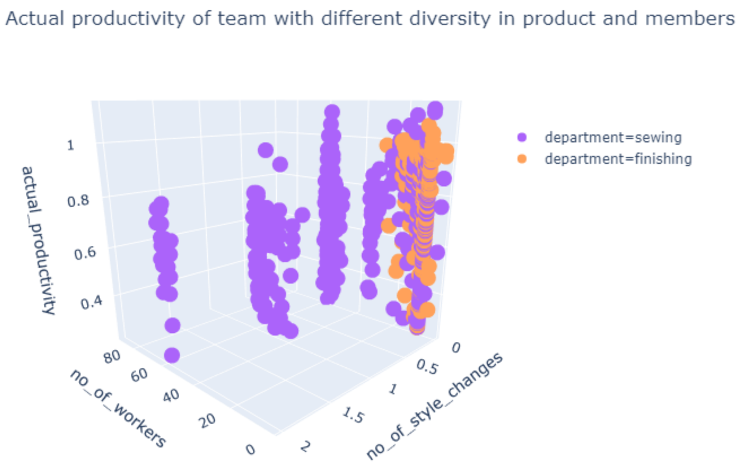


Figure 7: 3D (3 Dimensional) map between style changes, number of workers and actual productivity

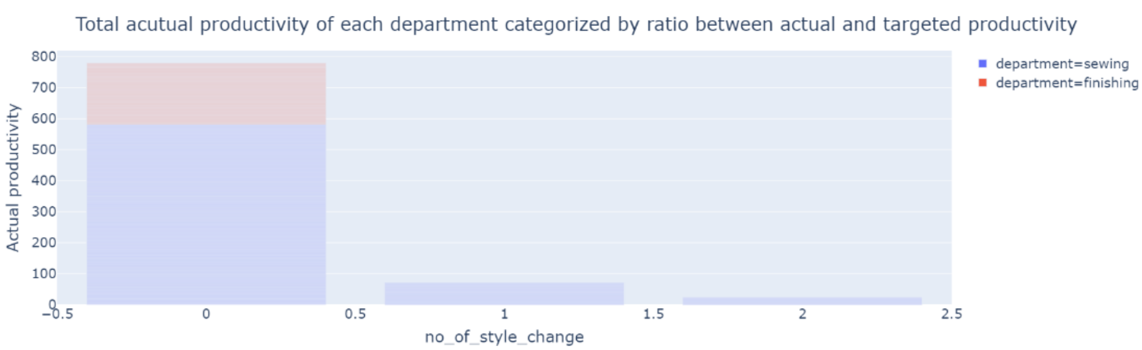


Figure 8: Styles changed quantities by departments.

Overall, the dataset shows not many changes in the style of all records. At only around 10% of the whole dataset has styles changed in the range of 0 to 2. The higher the number of styles changed, the lower actual productivity is achieved. This can be reflected from figure 7 that the highest actual productivity values at style changes 0,1,2 are 1.1, 0.97, 0.82, respectively. Moreover, the more changes of styles during garment work, the higher number of workers is required in the production team. As shown in the PPGE, more than 50 workers are demanded for the production process with only 2 changes of style. Therefore, from the given pattern, it can be concluded that the change in the style of the product affects actual productivity as well as consumes more labor force that could result in more financial consideration in production.

From figure 3, there is a correlation between style changes in each task and smv, idle_men. We could conclude that as the more change in the style of the product in the task, the team requires more time to complete which reduce the overall productivity of the team. There is also the increase in idle_men and idle_time number which means style changes reduce the efficiency of the production due to not exploiting the working time of the workers.

Question 5: How to make prediction and improvement if the pattern of the actual productivity is not evenly distributed for the declining trend of monthly productivity?

Besides machine learning models, multiple linear regression can also be applied as a conventional method of predicting the actual performance of garment workers. Despite having a prediction of much lower quality, the statistical parameters with a confidence interval of 95% consolidate our previous findings and pave the way for suggestions on how to improve the actual productivity in the future. The process started with data filtering where suspected outliers were removed from the original dataset based on box plots. In this case, only outliers of the targeted productivity were removed due to their low occurrences and thin scattering in the respective box plot. The outlier removal resulted in a filtered dataset with 1118 instances while the separated outlier data was chosen to be a test sample with 79 instances.

Feature selection was the next step providing essential correlation coefficients as well as the R squared values representing the amount of variation caused to workers' productivity (i.e. the response). Based on the heat map shown in figure 2, four input features were found to have noticeable correlations with the working efficiency of a group of workers on their allocated task, including their team number, targeted productivity, the number of idle men, and style changes required for the product. The subsequent residual analysis after the regression indicates a residual distribution acceptably skewed to the left with an approximate mean of 0. As a result, there is a variation of 20.09% in the response due to the identified features, while the F test has shown an F value of 70 compared to the critical value of 2.37 for k1 = 4 and k2 = 1113 at the confidence level of 0.05. This confirms the hypothesis that there is some relationship between at least one of the inputs and the response. The regression model is:

$$\hat{y} = 0.045854 - 0.00752x_{team} + 1.00911x_{target} - 0.00754x_{idle\ men} - 0.02514x_{style}$$

Another method to evaluate the model accuracy apart from the requirement of the R squared value above 50% is to define a baseline score as the Mean Absolute Error (MAE) between the mean of the actual productivity in the filtered data and the test data, meaning that the mean value is utilized as a rough prediction for the test sample:

$$baseline\ score = \frac{1}{N_{test}} \sum_{i=1}^{N_{test}} \left| y_{i,test} - \bar{Y}_{filtered} \right|$$

The baseline score was calculated as 0.2534. To prove its validation, the regression model's prediction must have an MAE less than the baseline score in the test sample. The MAE is formulated by:

$$MAE = \frac{1}{N_{test}} \sum_{i=1}^{N_{test}} \left| y_{i,test} - y_{i,predicted} \right|$$

In practice, the more viable prediction tool often used is machine learning. Not only can this type of models generate more accurate results, but it is also much more robust than linear regression. In the scope of this research, several machine learning models were built to find out which one is best used to predict the actual productivity and thus improvements can be made based on the predicted results.

Feature selection:

RFE (Recursive feature elimination) is used to eliminate features from the training dataset by searching and weighting recursively for a subset of features. At each stage of the search, we use an estimator to rank the features from the top priority to the lowest. The least important features are recursively discarded to rebuild the model. As a result, we archive a certain number of features for the prediction model.

RMSE value:

```
import math
X_train, X_test, y_train, y_test = train_test_split(rfe.transform(X), y, test_size=0.2, random_state=112)

model = RandomForestRegressor().fit(X_train, y_train)

y_pred = model.predict(X_test)
mse_value = mse(y_true=y_test, y_pred=y_pred)
#print(mse_value)
rmse = math.sqrt(mse_value)
print(f'RMSE = ',rmse)
```

```
RMSE =  0.1242877319460136
```

Feature results:

Features	Rankings
team	1
targeted_productivity	2
smv	3
over_time	4
incentive	5
no_of_workers	6

The top 6 values above generated from the RFE feature selection algorithm with **RMSE score = 0.124.** The results show a different perspective of the most impactful factor on the actual performance, which is the workers' team allocation compared to the assumption of targeted values as shown in the regression model. This opens up the possibility of leveraging effective coordination in the workplace as one of the measures to improve work efficiency.

Built models to predict the actual productivity:

We split and drop the four columns {quarter, department, day, team} into new columns using Label Encoding techniques. Then we use the 80/20 training/test split strategy on this dataset. Six models are generated including Logistic Regression, Decision Tree Classifiers, Random Forest, Support Vector Machine, k-Nearest Neighbor & Linear Discriminant Analysis. The results are as follows (The specific code explained for these models is included in the submission):

```
df_perf_metrics
```

	Model	Accuracy_Training_Set	Accuracy_Test_Set	Precision	Recall	f1_score
0	LogisticRegression	0.662143	0.665714	0.701299	0.603352	0.648649
1	DecisionTreeClassifier	0.998571	0.817143	0.848485	0.782123	0.813953
2	RandomForestClassifier	0.998571	0.882857	0.892045	0.877095	0.884507
3	SVC	0.666429	0.665714	0.706667	0.592179	0.644377
4	KNeighborsClassifier	0.845000	0.794286	0.836478	0.743017	0.786982
5	LinearDiscriminantAnalysis	0.785714	0.788571	0.766497	0.843575	0.803191
6	LinearDiscriminantAnalysis	0.785714	0.788571	0.766497	0.843575	0.803191

Table 2: Machine learning classifier validation & training results

From the result of different methods, the SCV shows overfit in training and test set, which cannot be used for prediction compared to the others. Logistic Regression model mitigate the overfitting; however, its accuracy does not satisfy compare with other models. The Random Forest Classifier model has overfitting ~ 10% but its accuracy is relatively high. Therefore, RFC is the best approach for this garment productivity dataset based on the highest accuracy score in the test set as well as the best overall performance.

5. Key findings and Suggestions:

The research questions addressed above aim to provide bases to derive a strategy for the improvement of the actual productivity. In practice, the authorities will set targeted values against each team to meet the production goals in time [5]. If these values are not properly allocated, the employees may have difficulties in meeting their work goals, resulting in undesired productivity gaps. The higher the gaps, the more likely the company faces a huge production loss. Therefore, a prediction of the actual productivity can hint the manufacturers at more accurate targets for the workers. This in turn minimizes the chance of production loss and maximizes profit.

Through the process of generating a multiple linear regression model for the productivity prediction, the statistical analysis has revealed not just the hidden interaction between various input factors and the response (i.e., the actual productivity) as well as among the inputs themselves, but also an established method for the derivation of similar prediction models in the future. In general, there were six main steps during the model building, comprising of data inspection, outlier removal, data splitting, feature selection, regression, and accuracy evaluation. The refined data served for the regression while the outliers split from the original dataset were utilized as a test sample. To evaluate the model accuracy, a baseline score was calculated as 0.2534, compared to the MAE of 0.1636 between the test sample (i.e., the outlier data) and the predicted productivity. Similarly, the MAE of the prediction in the original dataset is only 0.1071, hence the prediction model is statistically appropriate for further discussion.

Due to the proven bond between the targeted productivity and the actual one, this type of targets can be strategically set to reduce the productivity gap. Specifically, 14.54% of the variation in the actual productivity is due to the targeted values, showing the highest impact on the response among all other factors in the dataset. This shows an opposite trend to the initial observation where the working efficiency met the anticipation of 80% in only 250 out of 540 cases. The reason for such a discrepancy originates from the compensation of over-productive workers whose productivity ranges from 85% up to 110%, accounting for more than 300 instances in the record. In other words, the positive productivity gaps compensate the negative ones and thus increasing the correlation between the targeted and actual productivity. Consequently, the multiple regression model suggests that for each target assignment to a team, the members would normally meet 85.77% up to 116% of the value set by the associated industrial engineers with a 95% confidence interval. The estimated variation is, therefore, essential to the planning of the production parameters.

Another potentially influential input is the incentive paid, often believed to be a driving factor of higher productivity. The conducted test has shown that while it has some effect on the working efficiency, the incentive shares negligible correlations with the number of workers and overtime. For instance, even though team 9 was the most rewarded team during the 3-month timeframe, team 1 was the most productive, weakening the bond between the incentive and the workers' productivity. Also, whereas there was the least number of workers in March, the accumulated paid amount was surprisingly the highest, meaning that there was much more sewing required to be done for such a small group of employees and thus individuals might have to work overtime. Since the workers might normally have little interest in the incentive, they would have failed to meet the due time if there had been insufficient coworkers, which results in a proportion between the number of workers and over time rather than the involvement of any monetary reward. Based on the findings, ineffective work coordination comes up as a nuisance factor overriding the incentive paid and needs to be improved for better working efficiency.

Products' complexity also poses a great challenge to punctuality. In this regard, the number of style changes accounts for 2.85% of the variation in the response. A higher frequency of style change can reduce efficiency. In the dataset, since there is a chance of 12.28% that the staff must deal with a more complex garment, the actual productivity will accordingly decrease by 4% to 9% per style change, affecting the eventual profit. Such inflexibility in the garment market limits the variety of products due to the labour-intensive nature of the industry. Hence, with the emergence of more advanced technology, leadership vision is of prominence in making the garment industry more robust and in turn maximizing the productivity of employees as well as profit.

6. Conclusion:

From the first procedure of data cleaning to posing the hypothesis questions from the evaluation on the dataset, the initial process showed some valuable approaches to work on our data to later deprive our methodology to give the results of correlations. After performance data analysis, some key points are drawn out to observe the worker's daily performance to attributes from the dataset including targeted productivity incentive, idle factors, SMV, etc., for which the actual productivity is mainly under influence of the assigned productivity from the Authority. Using multiple regression, along with the application of Machine Learning, a prediction model is generated to make further implications that also shows relatively decent performance from the presented model. The prediction and suggestions, therefore, are prompted from these achievements and promisingly could be in actual implementation in the garment industry.

7. Reference:

[1] N. Kabeer and S. Mahmud, "Rags, Riches and Women Workers: Export-oriented Garment Manufacturing in Bangladesh", pp. 133-164, 2006. Available: https://www.wiego.org/sites/default/files/publications/files/Kabeer-Mahmud-Export-Oriented-Garment-Bangladesh.pdf?fbclid=IwAR3M1g9i_VFsnQEZqBTRIxxiwPorv9lyRPrQeT8evDZG-h4ER-AqspCU4RI. [Accessed 23 May 2021].

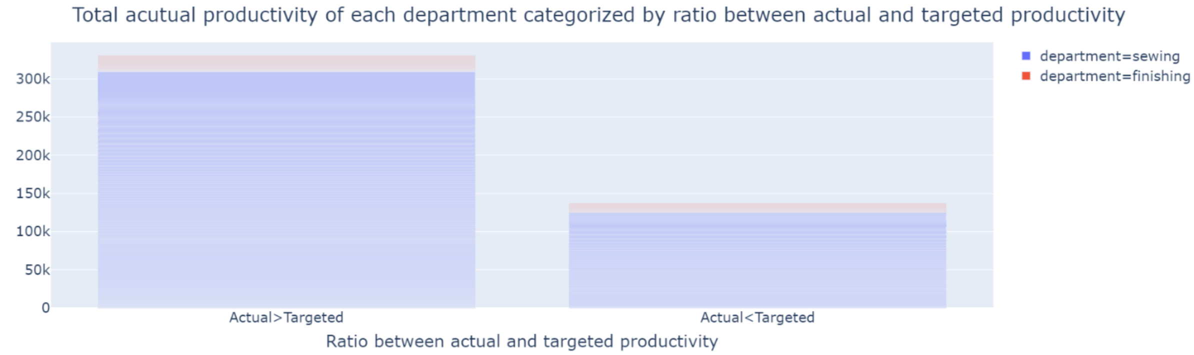
[2] A. Al Imran, "UCI Machine Learning Repository: Productivity Prediction of Garment Employees Data Set", *Archive.ics.uci.edu*, 2020. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/Productivity+Prediction+of+Garment+Employees. [Accessed: 27- May- 2021].

[3] C. Hossan, M. Rahman Sarker and R. Afroze, "Recent Unrest in the RMG Sector of Bangladesh: Is this an Outcome of Poor Labour Practices?", *International Journal of Business and Management*, vol. 7, no. 3, 2012.

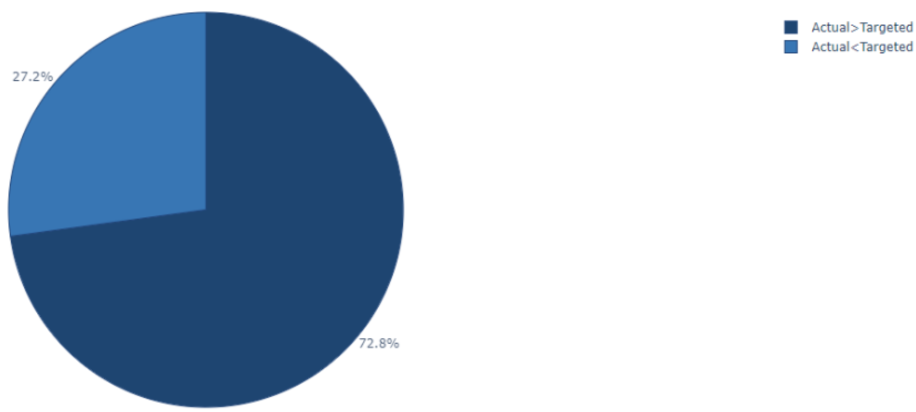
[4] P. Saha and S. Mazumder, "Impact of Working Environment on Less Productivity in RMG Industries: A Study on Bangladesh RMG Sector", Global Journals Inc. (USA), vol. 15, no. 2, 2021.

[5] A. A. Imran, M. N. Amin, M. R. Islam Rifat, and S. Mehreen, “Deep Neural Network Approach for Predicting the Productivity of Garment Employees”, 2019.

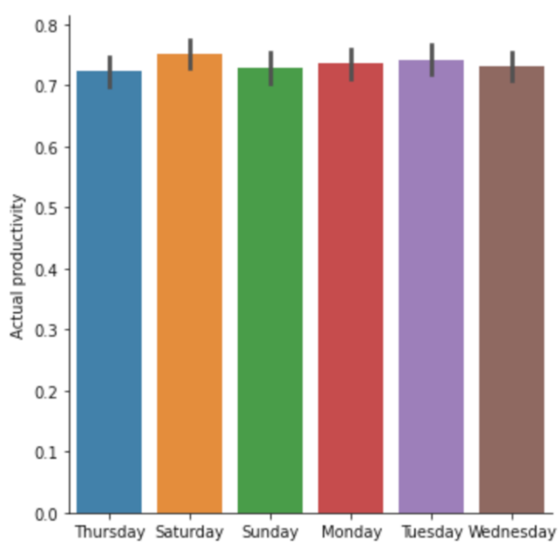
8. Appendix:

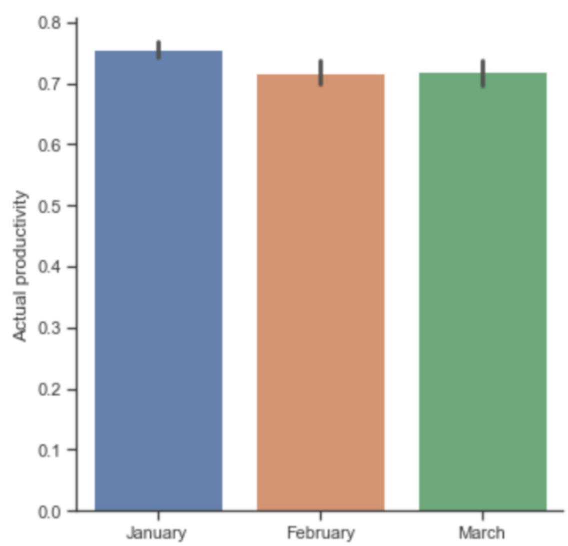


Appendix 1: Ratio between Actual and Targeted productivity by departments

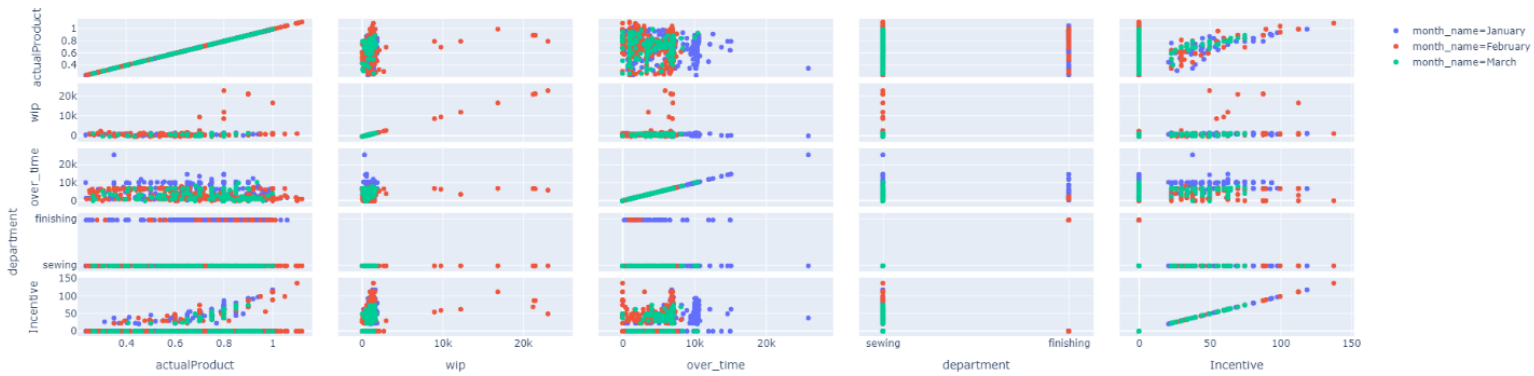


Appendix 2: Ratio between actual and Targeted productivity probability

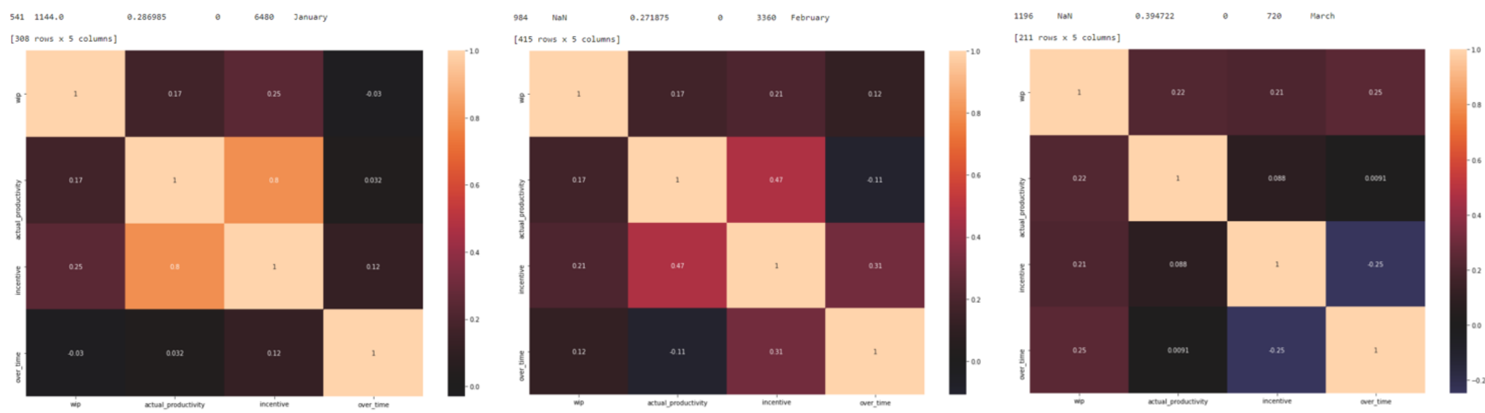




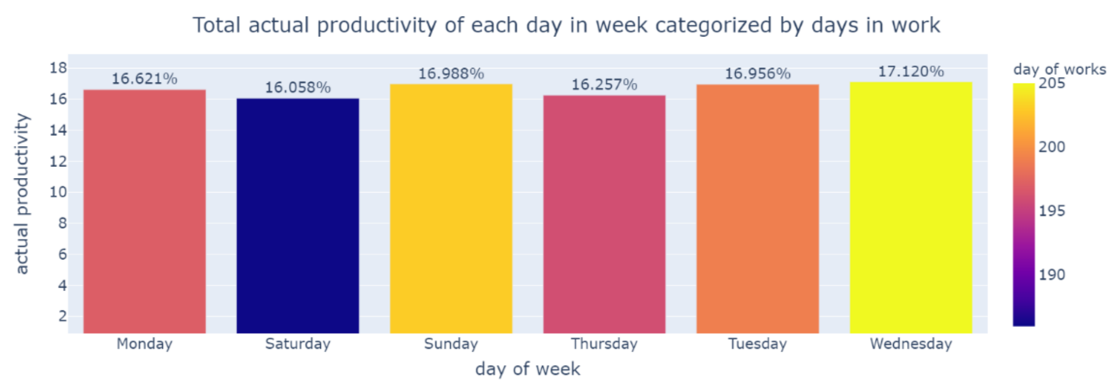
Appendix 3. Actual productivity in weekdays(left) and months.



Appendix 4: 3D map relationships between the incentive to actual productivity and WIP by departments.



Appendix 5: Heatmap of correlation between WIP, actual productivity, incentive and overtime in January (left), February (middle) and March.



Appendix 6: Actual productivity recorded in weekdays.

Appendix 7: Does the actual productivity of each team fluctuate to certain patterns? What are the possible factors behind those patterns? Which one is the most impactful?

In terms of attributes correlation shown in the heatmap of Figure 3, the strongest bond is the number of workers - overtime sharing 0.73. (Standard Minute Value – overtime), which stays true to the working time of each worker extending accordingly to the estimated time spending on a particular task.

The actual productivity, contrary to the initial prediction, is not related to the number of workers in a day, for both showing an extremely low match.

Evaluating the relations between actual productivity and number of styles change, idle men, idle time, the recorded values respectively are -0.21, -0.18, -0.081, showing all as negative, which means these parameters are inversely proportional to the actual performance of garment workers.

As in the factual scenario, the productivity is of major dependence on the workers' labor, therefore, there should be as little idle men and idle time as possible. Similarly, in the case of style changes, more accomplishments are produced if the process requires the least changes in garment designs, as the shifting phases would take up the efficient time of working.