Rohitashwa Chakraborty,   
Divyansh Karki,   
Safiuddin Mohammed,   
Yen Wen Ting

Project goals:

The garment industry is a primary example of industrial globalization. It is highly labour-intensive and with many manual processes still involved. Therefore the to meet the demand of the industry, it is crucial for decision-makers to analyse and predict the productivity of their workers About the Data :

* Source: Kaggle.
* Output Variable :

|  |  |
| --- | --- |
| actual productivity | Measured actual productivity |

* Possible Tasks:

1. Impact of different variables on the productivity of workers. For Example:
   1. Distribution of productivity over different quarters of the month.
   2. Which department seems to be the most productive
   3. Does the team size increase or decrease productivity?
   4. Is incentive the primary motivator towards productivity of workers?
2. Regression / Classification on the dataset:
   1. Apply the various machine learning models, to predict the productivity of workers.

We could potentially apply a model that could categorize the “environment” as highly productive / moderately productive / low productivity.

**Exploratory Data Analysis:**

Ours was a publicly available dataset on the productivity of daily wage workers in a garment factory[]. The dataset has team level data for two of its departments: Finishing and Sewing. The *Finishing* department didn’t have any *wip* data and the latter was dropped during analysis. Team level metrics like *overtime, incentive* etc. were scaled down to per hour-per member which were more comparable. Furthermore, the two departments have significant parity in terms of nature of work. There exists large difference in the skill and complexity of the tasks and thus, the dataset was split, and the departments were analysed individually. This allowed us to study their trends are as follows:

Teams seem to have developed a specialization that makes them more productive in one department than the other[]. It was also observed that in the both the departments, productivity is NOT a proxy for consistency. There are teams which exceed their overall target productivity over the first quarter by about 17% but are very volatile. On the other hand, there are teams that might have underperformed but are highly reliable in terms of their turnover. The aim of the EDA is to determine the factors which improve productivity. There wasn’t enough data to analyse performance consistency.

*Team Size v/s Productivity:* While “Finishing” seems to favor larger teams, “Sewing” experiences a sharp drop in productivity as team size increases. This could be explained by the nature of work. Perhaps the tasks in the former are more independent and thus larger teams directly increase productivity. However, in the case of sewing, the tasks might involve more team work and coordination, which would explain the presence of an ideal team size, which when exceeded, drastically hampers productivity.

*Task Complexity v/s Productivity:* It was observed that teams in “Finishing” department typically have highest productivity when the task typically takes 3-6 minutes while teams in “Sewing” are most productive in the 23–33-minute block. This implies that “Finishing” jobs are typically less complex and, from our last discussion- largely independent, thus, can be organised in larger groups than “Sewing” jobs which are more specialized.

*Overtime:* Teams in “Finishing” displayed no observable improvement in productivity with incentives, while “Sewing” typically witnesses higher productivity rates when made to work overtime for more than 2.5 hours.

*Incentive:* As expected, both departments show a high positive correlation between their Incentive and productivity.

Appendix: