**Dataset link:** <https://archive.ics.uci.edu/dataset/597/productivity+prediction+of+garment+employees>

#### About the dataset:

This dataset includes important attributes of the garment manufacturing process and the productivity of the employees which had been collected manually and also been validated by the industry experts.

Satisfying the huge global demand for garment products is mostly dependent on the production and delivery performance of the employees in the garment manufacturing companies So, it is highly desirable among the decision makers in the garments industry to track, analyse and predict the productivity performance of the working teams in their factories.

#### Variables details:

| **Variable Name** | **Role** | **Type** | **Description** | **Units** | **Missing Values** |
| --- | --- | --- | --- | --- | --- |
| date | Feature | Date |  |  | no |
| quarter | Feature | Categorical |  |  | no |
| department | Feature | Categorical |  |  | no |
| day | Feature | Categorical |  |  | no |
| team | Feature | Integer |  |  | no |
| targeted\_productivity | Feature | Continuous |  |  | no |
| smv | Feature | Continuous |  |  | no |
| wip | Feature | Integer |  |  | yes |
| over\_time | Feature | Integer |  |  | no |
| incentive | Feature | Integer |  | BDT | no |
| idle\_time | Feature | Integer |  |  | no |
| idle\_men | Feature | Integer |  |  | no |
| no\_of\_style\_change | Feature | Integer |  |  | no |
| no\_of\_workers | Feature | Integer |  |  | no |
| actual\_productivity | Target | Continuous |  |  | no |

#### Additional Variable Information

01 date : Date in MM-DD-YYYY

02 day : Day of the Week

03 quarter : A portion of the month. A month was divided into four quarters

04 department : Associated department with the instance

05 team\_no : Associated team number with the instance

06 no\_of\_workers : Number of workers in each team

07 no\_of\_style\_change : Number of changes in the style of a particular product

08 targeted\_productivity : Targeted productivity set by the Authority for each team for each day.

09 smv : Standard Minute Value, it is the allocated time for a task

10 wip : Work in progress. Includes the number of unfinished items for products

11 over\_time : Represents the amount of overtime by each team in minutes

12 incentive : Represents the amount of financial incentive (in BDT) that enables or motivates a particular course of action.

13 idle\_time : The amount of time when the production was interrupted due to several reasons

14 idle\_men : The number of workers who were idle due to production interruption

15 actual\_productivity : The actual % of productivity that was delivered by the workers. It ranges from 0-1.

#### Observations:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| date | department | day | month | team\_no | targeted\_productivity | actual\_productivity |
| 01-01-2015 | finishing | Thursday | January | 1 | 75.00% | 88.65% |
| 01-01-2015 | sweing | Thursday | January | 1 | 75.00% | 75.04% |

1. Each row represent A unique team productivity score in one day for particular department(team\_id)
2. Team members count in each group is not clear and constant, there is some variabilty. As the scores are focusing more on the teams not individual we will also do analysis based on teams scores

#### Data Cleaning Actions

1. Corrected the Date column datatype.
2. Ensured four quarters per month as specified.
3. Added Month column and Productivity boolean value.
4. Converted no\_of\_workers to integer.
5. Filled null WIP values with 0 (indicating no work).
6. Changed targeted and actual productivity to percentages.

#### Exploratory Data Analysis (EDA)

1. **Performance Comparison**:
   * Utilized basic statistics, histograms, box plots, and a Z tests
   * 73% of the time, teams met their targets; 23% failed.
   * A one-tailed Z- test showed an evidance for higher actual mean productivity score than target at 95% confidence level.
   * Average targeted productivity: 72.96% (low variability, SD = 9.7%)
   * Average actual productivity: 73.5% (high variability, SD = 17.4%)
2. **Productivity by Quarter and Day**:
   * Highest productivity in Q1, lowest in Q3 (limited data for Q3).
   * Significant difference in productivity scores across months (ANOVA test).
   * Higher productivity on Saturdays, Mondays, and Tuesdays.

#### Dashboard Development

Stakeholders requested a dashboard to monitor team performance:

1. Compare actual vs. targeted average productivity for each team.
2. Analyze overtime by each team to identify issues and optimize processes.
3. Examine the relationship between idle time, idle men, and overtime.
4. Include slicers for Month, Quarter, and Day to filter performance data accordingly.

This comprehensive analysis and visualization aim to enhance productivity tracking and decision-making in garment manufacturing.