Productivity Estimation Model for Warehouse Operations

Hack MTY 2025 Submission

Your Team Name

October 25, 2025

Abstract

This report details a linear regression model developed for the Hack MTY 2025 hackathon to estimate the time required for packing warehouse drawers and trolleys. By analyzing historical efficiency data from 120 records covering 1,717 items, the model predicts P50 (median) and P90 (90th percentile) packing times. These predictions are then compared against operational targets (P50: 5.0 minutes, P90: 7.0 minutes) to perform a gap analysis, providing key performance indicators (KPIs) for warehouse productivity. The model demonstrates strong performance, meeting the P50 target with a margin of -0.69 minutes while slightly exceeding the P90 target by +0.21 minutes, highlighting opportunities for variability reduction.

Contents

1 Introduction

Efficient warehouse operations are critical for supply chain performance and customer satisfaction. A key process in modern warehousing is the packing of items into drawers, which are then consolidated into trolleys for distribution. Accurately estimating the time required for this task is essential for resource planning, setting performance targets, and identifying operational bottlenecks.

This project, developed for Hack MTY 2025, introduces a data-driven measurement engine to predict packing times. The primary objectives are:

- To develop a simple, interpretable model based on historical operational data.

- To predict both the median (P50) and a high-percentile (P90) completion time to account for operational variability.

- To provide a quantitative gap analysis by comparing predictions against established operational KPIs.

- To deliver actionable insights for operational planning and performance management.

The model serves as both a predictive tool and a diagnostic instrument, enabling warehouse managers to understand current performance levels and identify areas for improvement.

2 Methodology

2.1 Data Source

The model was trained on a dataset containing historical employee efficiency records (`employee_efficiency.csv`) and productivity estimation data (`productivity_estimation.csv`). The efficiency dataset comprises 120 individual packing task records, covering a total of 1,717 items packed over 6,348 seconds of operational time. Each record includes:

- N_{items}: The number of items packed in a drawer

- T_{drawer}: The duration of the packing task in seconds

The productivity dataset contains 100 drawer configurations used for prediction and validation purposes.

2.2 Mathematical Model

We propose a linear model to describe the relationship between the number of items and the packing time. The packing time for a single drawer, T_{drawer}, is modeled as:

$$T_{\text{drawer}} = t_0 + s \cdot N_{\text{items}} \tag{1}$$

where:

- t_0 is the fixed setup time per drawer (in seconds)

- s is the average time required to pack a single item (in seconds per item)

- N_{items} is the number of items in the drawer

For this implementation, the setup time t_0 was set to zero based on initial analysis, simplifying the model to a direct proportionality. The parameter s was estimated from the historical data by calculating the ratio of total duration to total items packed:

$$\hat{s} = \frac{\sum_{i=1}^{n} T_i}{\sum_{i=1}^{n} N_i} \tag{2}$$

2.3 P50 and P90 Prediction

To account for operational variability and provide robust predictions, we estimate two key metrics:

- **P50 Time (Median)**: The expected time, representing typical performance. This is calculated directly from the linear model:

$$T_{\text{P50}}(N) = t_0 + \hat{s} \cdot N_{\text{items}} \tag{3}$$

- **P90 Time (90th Percentile)**: A conservative estimate accounting for variability and delays. This is calculated by adding the 90th percentile of the model's absolute residuals (q_{90}) to the P50 time:

$$T_{\text{P90}}(N) = T_{\text{P50}}(N) + q_{90,\text{residual}} \tag{4}$$

The residuals represent the difference between observed and predicted times, capturing the natural variability in the packing process. The 90th percentile of absolute residuals provides a robust measure of this variability.

2.4 Trolley-Level Aggregation

For a trolley containing multiple drawers, the total predicted time is the sum of individual drawer predictions:

$$T_{\text{trolley}} = \sum_{j=1}^{D} T_{\text{drawer},j} \tag{5}$$

where D is the number of drawers in the trolley.

3 Results

3.1 Model Parameters

The model was fitted using the efficiency dataset. The estimated parameters are summarized in Table 1.

Table 1: Estimated Model Parameters from Historical Data

Parameter	Value	Unit
Setup time (t_0)	0.00	seconds
Seconds per item ($\hat{s}$)	3.70	seconds/item
P90 residual (q_{90})	34.79	seconds
Mean absolute residual	18.72	seconds
Standard deviation of residuals	22.10	seconds

Estimated from 120 efficiency records totaling 1,717 items and 6,348 seconds.
The model explains packing time as approximately 3.70 seconds per item with a P90 variability buffer of 34.79 seconds.

The parameter $\hat{s} = 3.6971$ seconds/item indicates that, on average, each item requires approximately 3.7 seconds to pack. The P90 residual of 34.79 seconds represents the additional time buffer needed to account for 90% of observed variability in the process.

3.2 Data Summary

Table 2 provides an overview of the datasets used for model training and prediction.

Table 2: Dataset Summary Statistics

Metric	Value
Productivity data records	100 drawers
Efficiency data records	120 records
Total items (efficiency dataset)	1,717 items
Total duration (efficiency dataset)	6,348 seconds
Average items per record	14.3 items
Average duration per record	52.9 seconds

3.3 Prediction Examples

Table 3 presents the model's predictions for a representative drawer and trolley configuration.

Table 3: Sample Predictions for Drawer and Trolley Configurations

Entity	Items	P50 (min)	P90 (min)	Variability
Sample Drawer	12.00	0.74	1.32	0.78
Sample Trolley (5 drawers)	70.00	4.31	7.21	0.67

Variability is calculated as $(T_{P90} - T_{P50})/T_{P50}$.
The sample trolley comprises 5 drawers with a total of 70 items.

The variability metric indicates the relative uncertainty in predictions. The drawer-level variability of 78.4% is higher than the trolley-level variability of 67.2%, demonstrating that aggregation across multiple drawers reduces relative uncertainty.

4 Analysis

4.1 KPI Gap Analysis

The model's predictions for the sample trolley were compared against operational targets. The operational KPIs specify a P50 completion time of 5.0 minutes and a P90 time of 7.0 minutes. Table 4 presents the gap analysis results.

Table 4: KPI Gap Analysis for Sample Trolley (70 items, 5 drawers)

Metric	Target (min)	Predicted (min)	Gap (min)	Status
P50 Time	5.00	4.31	−0.69	MEETS TARGET
P90 Time	7.00	7.21	0.21	EXCEEDS TARGET

Negative gap indicates predicted time is below target (better performance).
Positive gap indicates predicted time exceeds target (performance issue).

4.2 Discussion

The gap analysis reveals important insights about current operational performance:

- **P50 Performance**: The model predicts a median packing time of 4.31 minutes, which is 0.69 minutes (14%) below the target of 5.0 minutes. This indicates that typical performance meets and exceeds expectations, suggesting well-trained staff and efficient processes under normal conditions.

- **P90 Performance**: The 90th percentile prediction of 7.21 minutes exceeds the target of 7.0 minutes by 0.21 minutes (3%). While this gap is relatively small, it indicates that the process occasionally experiences variability that pushes completion times beyond acceptable thresholds. This suggests opportunities for:

  - Process standardization to reduce variability
  - Investigation of outlier cases contributing to delays
  - Training interventions for consistency improvement
  - Potential workflow optimizations to reduce setup or handling time

- **Variability Impact**: The 67.2% variability at the trolley level means that the P90 time is approximately 67% longer than the P50 time. This substantial variability is the primary driver of the P90 gap and represents the most significant opportunity for operational improvement.

5 Validation and Acceptance

5.1 Model Validation Criteria

The model was subjected to a series of acceptance tests to ensure its validity and reliability. Table 5 summarizes the validation results.

Table 5: Model Validation and Acceptance Criteria

Validation Check	Criterion	Result
Non-negative setup time	$t_0 \geq 0$	PASS
Positive item time	$\hat{s} > 0$	PASS
Positive residual variability	$q_{90} > 0$	PASS
Predictions scale with items	$T_{\mathrm{P90,trolley}} > T_{\mathrm{P90,drawer}}$	PASS
Measurable KPI gap	$\mathrm{Gap} \neq 0$	PASS

All validation criteria were successfully met, confirming the model's mathematical soundness and predictive validity.

5.2 Statistical Diagnostics

The model demonstrates reasonable fit quality:

- **Mean Absolute Residual**: 18.72 seconds indicates typical prediction error

- **Standard Deviation**: 22.10 seconds captures the spread of residuals

- **P90 Residual**: 34.79 seconds provides a conservative buffer for 90% of cases

The residual statistics suggest that while the linear model captures the general trend, there is inherent process variability that cannot be explained by item count alone. Future model enhancements could incorporate additional features such as item type, employee experience, time of day, or drawer configuration complexity.

6   Conclusion

This report presented a linear regression-based measurement engine for predicting warehouse packing times at both drawer and trolley levels. The model successfully quantifies the relationship between the number of items and packing duration, providing valuable P50 (median) and P90 (90th percentile) estimates.

Key findings include:

1. The model achieves an estimated packing rate of 3.70 seconds per item with a P90 variability buffer of 34.79 seconds.

2. Median performance (P50) exceeds operational targets by 0.69 minutes, indicating strong baseline productivity.

3. High-percentile performance (P90) slightly exceeds targets by 0.21 minutes, highlighting variability as the primary improvement opportunity.

4. The model demonstrates mathematical validity and passes all acceptance criteria.

6.1   Recommendations

Based on the analysis, we recommend the following actions:

- **Variability Reduction**: Investigate root causes of process variability, particularly cases in the upper percentiles. This could involve process standardization, ergonomic improvements, or workflow optimization.

- **Continuous Monitoring**: Deploy the model as a real-time monitoring tool to track performance against targets and detect degradation early.

- **Feature Enhancement**: Incorporate additional predictive features such as item type, employee experience level, shift timing, and drawer complexity to improve prediction accuracy.

- **Model Refinement**: Re-estimate the setup time parameter (t_0) as a learnable parameter rather than fixing it to zero, which may improve fit quality.

- **Outlier Analysis**: Conduct detailed analysis of records in the upper 10th percentile to identify specific factors contributing to delays.

6.2   Future Work

Potential extensions of this work include:

- Exploring non-linear models (e.g., polynomial regression, splines) to capture potential saturation effects at high item counts.

- Developing employee-specific models to enable personalized productivity tracking and targeted training.

- Integrating the model into warehouse management systems for real-time operational decision support.

- Expanding the model to predict energy consumption, ergonomic stress, or quality metrics in addition to time.

- Building predictive models for other warehouse operations (picking, sorting, loading) using similar methodologies.

The measurement engine developed for this project demonstrates the value of data-driven approaches to operational performance management and provides a foundation for continuous improvement in warehouse productivity.