**DATA ANALYTICS CASE STUDY**

**FMCG Warehouse: Optimizing Amazon's Distribution Efficiency**



**Presented BY**

**JAYA SUHANE| JOB AAJ LEARNINGS |24/03/2025**

# Problem

Amazon is experiencing inefficiencies in its warehouse operations, which are affecting the supply chain, leading to stockouts, delivery delays, and increased costs. Your task is to analyse warehouse data to uncover the root causes of these issues and propose data-driven solutions to optimize warehouse performance.

# Background

You are a data analyst at Amazon, which has an extensive network of warehouses across various regions. The company is facing challenges in ensuring efficient warehouse operations, resulting in frequent stockouts, delays in deliveries, and increased operational costs. The management has tasked you with analysing the data from these warehouses to identify key issues and provide actionable insights to enhance warehouse efficiency.

# Solution

To solve Amazon's warehouse inefficiencies as a data analyst, we can take a systematic approach involving data exploration, root cause analysis, and optimization strategies.

**1. Understand the Problem & Define Key Metrics**

**Pain Points:**

- Stockouts → Lost sales & customer dissatisfaction

- Delivery delays → Reduced Prime efficiency & SLA breaches

- Increased costs → Inefficient labor & warehouse space usage

**Key Performance Indicators (KPIs):**

- **Order Fulfilment Rate** = (Orders shipped on time / Total orders) * 100

- **Inventory Turnover** = Cost of Goods Sold (COGS) / Average Inventory

- **Warehouse Capacity Utilization** = (Used storage space / Total storage space) * 100

- **Picking Accuracy Rate** = (Correctly picked orders / Total orders) * 100

- **Average Order Processing Time** = Time taken from order placement to dispatch

- **Dock-to-Stock Time** = Time taken to unload and store new inventory

---

## 2. Collect & Analyse Warehouse Data

**Data Sources:**

- Warehouse Management System (WMS) logs

- RFID & IoT sensor data for real-time tracking

- Labor productivity records

- Historical order & shipment data

- Inventory records & demand forecasts

**Techniques for Analysis:**

- **Exploratory Data Analysis (EDA):** Identify trends, outliers, and correlations

- **Time Series Analysis:** Detect seasonal demand fluctuations

- **Process Mining:** Map warehouse workflow bottlenecks

- **Heatmaps:** Visualize inefficient storage areas & high-traffic zones

- **Predictive Analytics:** Forecast demand and prevent stockouts

- **Cluster Analysis:** Identify fast-moving vs. slow-moving inventory

# 3. Identify Root Causes of Inefficiencies

- 🔍 **Potential Issues & Data Insights**

| Issue | Data Indicators | Possible Solution |
|---|---|---|
| **Stockouts** | High demand volatility, poor forecasting | AI-driven demand prediction |
| **Delivery Delays** | Long pick & pack times, inefficient layout | Optimize warehouse zoning & automation |
| **Inventory Inaccuracy** | Mismatch between WMS and physical stock | RFID-based real-time tracking |
| **Poor Labor Productivity** | High idle time, overtime costs | Workforce optimization & task scheduling |
| **Underutilized Space** | Low storage density, poor slotting | Dynamic slotting & vertical storage |

## 4. Implement Data-Driven Optimization Strategies

◈ Improve Inventory Management:

- Demand Forecasting Models → ML-based demand prediction

- ABC Analysis → Prioritize fast-moving SKUs near dispatch zones

- Safety Stock Optimization → Adjust buffer stock dynamically

◈ Optimize Warehouse Layout & Storage:

- Zone Picking → Reduce walking distance for pickers

- Automated Storage & Retrieval Systems (ASRS) → Faster access to items

- Heatmap Analysis → Relocate high-demand items near exits

◈ Enhance Order Fulfilment Process:

- Task Assignment via AI → Dynamic workforce allocation

- Barcode/RFID Tracking → Reduce picking errors

- Automated Packaging & Sorting → Improve throughput

◈ Improve Workforce Productivity:

- Labor Performance Dashboards → Monitor efficiency

- Predictive Workforce Scheduling → Adjust shifts based on demand

- Gamification & Incentives → Motivate warehouse workers

◈ Use IoT & Robotics for Automation:

- Robotic Process Automation (RPA) → Reduce manual work

- AI-driven AGVs & Conveyor Systems → Improve material flow

---

5. Continuous Monitoring & Optimization

- Real-time KPI Dashboards (e.g., Power BI, Looker Studio)

- A/B Testing for Layout & Process Changes

- Monthly Review of Performance Metrics

- Automated Alerts for Stockouts & Delays

## Project Scope And Methodology

### 🔨 Project Scope

This project aims to analyse Amazon's warehouse operations to identify inefficiencies contributing to stockouts, delivery delays, and increased costs. The focus will be on diagnosing operational bottlenecks, leveraging data analytics for insights, and proposing data-driven solutions to optimize performance.

### ✅ Objectives:

1. **Identify key inefficiencies** in inventory management, order fulfillment, and warehouse operations.

2. **Analyse warehouse data** to uncover root causes of stockouts, delays, and cost escalations.

3. **Develop predictive models & optimization strategies** to enhance warehouse efficiency.

4. **Implement data visualization** to monitor KPIs and drive decision-making.

## 🔍 Key Focus Areas:

- **Inventory Management**: Stock replenishment, demand forecasting, inventory accuracy.

- **Order Fulfilment**: Picking, packing, sorting, and shipment processing times.

- **Warehouse Layout & Storage**: Space utilization, item placement, and traffic flow.

- **Labor Productivity**: Workforce allocation, task scheduling, and operational efficiency.

- **Technology & Automation**: RFID tracking, robotics, and real-time monitoring.

## ⚒️ Methodology

**Phase 1: Problem Definition & Data Collection**

### ◈ Understanding the Business Challenge

- Conduct stakeholder interviews (warehouse managers, logistics teams).

- Identify major pain points: stockouts, delays, cost overruns.

- Define key performance indicators (KPIs).

### ◈ Data Collection

- Extract historical warehouse data from Amazon's Warehouse Management System (WMS).

- Collect data from IoT sensors, barcode/RFID scanners, and order logs.

- Integrate labor productivity records, storage utilization data, and shipment tracking.

**amazon**

📊 **Data Sources:**

| Data Type | Source | Metrics Collected |
|---|---|---|
| Inventory Data | WMS, ERP | Stock levels, stockout rates, replenishment frequency |
| Order Fulfillment | WMS, OMS | Processing time, picking accuracy, delivery times |
| Warehouse Layout | IoT, RFID, Sensors | Heatmaps, item placement, congestion zones |
| Labor Productivity | HR Systems, Work Logs | Worker efficiency, idle time, task completion rates |
| Cost Analysis | Financial Reports | Storage costs, labor costs, operational expenses |

**Phase 2: Exploratory Data Analysis (EDA) & Root Cause Analysis**

◈ **Data Cleaning & Preparation**

- Handle missing values, duplicate records, and inconsistencies.
- Standardize data formats across different sources.

◈ **Exploratory Data Analysis (EDA)**

- **Descriptive Analytics**: Identify trends, patterns, and outliers.
- **Correlation Analysis**: Find relationships between variables (e.g., stockouts vs. demand forecast errors).
- **Time Series Analysis**: Detect seasonality in demand and warehouse operations.
- **Process Mining**: Map workflows to identify inefficiencies in picking, packing, and shipping.

◈ **Root Cause Analysis (RCA)**

| Issue | Potential Cause | Data-Driven Insight |
|---|---|---|
| Stockouts | Poor demand forecasting | Forecast error rates, safety stock levels |
| Delivery Delays | Inefficient order picking & packing | High order processing times, congestion heatmaps |
| Increased Costs | Labor inefficiencies & underutilized space | Overtime costs, idle time analysis, storage density |

**Phase 3: Predictive Modelling & Optimization**

◈ **Demand Forecasting**

- Train **time-series forecasting models** (ARIMA, Boost, LSTM) on historical demand data.

- Optimize reorder points & safety stock levels to prevent stockouts.

◈ **Warehouse Layout Optimization**

- Use **clustering algorithms (K-Means, DBSCAN)** to group high-demand SKUs together.

- Implement **heatmap analysis** for optimal item placement.

◈ **Workforce & Order Fulfilment Optimization**

- Develop a **predictive labour scheduling model** based on peak demand hours.

- Implement **AI-driven picking route optimization** (e.g., shortest path algorithms).

- Recommend automation (robotic pickers, ASRS) for repetitive tasks.

**Phase 4: Implementation & Data Visualization**

◈ **Dashboard Development (Power BI / Looker Studio)**

- **Live KPI tracking** (order processing time, fulfillment rates, cost trends).

- **Real-time alerts** for low inventory, delays, and bottlenecks.

- **Geospatial Analysis** for warehouse congestion & layout optimization.

◈ **A/B Testing & Continuous Monitoring**

- Test proposed warehouse process improvements on a smaller scale.

- Monitor the impact using real-time dashboards.

- Iterate strategies based on results & stakeholder feedback.

## Goals And KPI's

🎯 **Project Goals**

The primary goal is to **optimize warehouse efficiency** by analysing data to identify bottlenecks, reduce costs, and improve order fulfilment.

✅ **Specific Goals:**

1. **Reduce Stockouts** → Improve inventory accuracy and demand forecasting.

2. **Minimize Delivery Delays** → Optimize order picking, packing, and dispatch.

3. **Decrease Operational Costs** → Enhance labor productivity and space utilization.

4. **Improve Warehouse Throughput** → Reduce order processing time and increase automation.

5. **Enhance Customer Satisfaction** → Faster and more accurate deliveries.

📊 **Key Performance Indicators (KPIs)**

To measure the success of warehouse optimization, we track **efficiency, accuracy, and cost-related KPIs**.

**amazon**

## 📦 Inventory Management KPIs

| KPI | Formula | Goal |
| --- | --- | --- |
| Stockout Rate | (Stockouts / Total SKUs) * 100 | **<5%** (Reduce missed sales) |
| Inventory Accuracy | (System Stock / Physical Stock) * 100 | **>98%** (Ensure correct inventory tracking) |
| Inventory Turnover | COGS / Average Inventory | **>8-10x/year** (Faster stock movement) |
| Dead Stock Percentage | (Unsold Inventory >90 days / Total Inventory) * 100 | **<5%** (Optimize stock levels) |
| Replenishment Cycle Time | Time from reorder to shelf availability | **Reduce by 30%** (Faster restocking) |

## 🚚 Order Fulfilment KPIs

| KPI | Formula | Goal |
| --- | --- | --- |
| Order Fulfilment Rate | (Orders shipped on time / Total orders) * 100 | **>98%** (Meet SLAs) |
| Order Processing Time | Time from order placement to shipment | **Reduce by 40%** (Faster order handling) |
| Picking Accuracy | (Correctly picked orders / Total orders) * 100 | **>99.5%** (Reduce returns due to wrong items) |
| Dock-to-Stock Time | Time from truck unloading to inventory storage | **Reduce by 25%** (Faster stock availability) |
| Backorder Rate | (Backordered items / Total items ordered) * 100 | **<2%** (Ensure in-stock availability) |

# amazon

## 🏬 Warehouse Operations KPIs

| KPI | Formula | Goal |
|---|---|---|
| **Warehouse Capacity Utilization** | (Used storage space / Total storage space) * 100 | **>90%** (Optimize space) |
| **Throughput Rate** | Total orders processed per hour | **Increase by 20%** (Improve efficiency) |
| **Order Cycle Time** | Time from order receipt to shipping | **Reduce by 35%** (Faster processing) |
| **Travel Time per Pick** | Average time spent by workers retrieving items | **Reduce by 30%** (Optimize storage layout) |
| **Automation Utilization Rate** | (Automated tasks / Total tasks) * 100 | **>60%** (Increase robotic picking efficiency) |

## 💰 Cost & Productivity KPIs

| KPI | Formula | Goal |
|---|---|---|
| **Cost per Order Fulfilment** | Total warehouse operating cost / Orders processed | **Reduce by 25%** (Lower expenses) |
| **Labor Productivity** | Orders fulfilled per worker per hour | **Increase by 30%** (Optimize workforce) |
| **Shrinkage Rate** | (Lost or damaged inventory / Total inventory) * 100 | **<1%** (Improve handling & security) |
| **Energy Consumption per Order** | Energy used / Orders processed | **Reduce by 20%** (Enhance sustainability) |
| **Return Rate Due to Warehouse Errors** | (Returns due to picking errors / Total orders) * 100 | **<1%** (Improve accuracy) |

## Concept Used

To analyse and optimize Amazon's warehouse operations, we apply various **data-driven methodologies, optimization techniques, and predictive analytics**. Below is the core concepts used:

---

### 1 Data Analytics & Business Intelligence (BI)

📌 **Concepts Used:**

- **Exploratory Data Analysis (EDA)** → Identifies patterns, trends, and outliers in warehouse data.

- **Descriptive Analytics** → Summarizes current warehouse performance through KPI tracking.

- **Data Visualization** → Power BI, Looker Studio, Tableau dashboards for real-time monitoring.

📊 **Example Application:**

- Analysing warehouse throughput using **heatmaps** to detect congestion in picking routes.

- Creating **real-time dashboards** to monitor **stock levels, fulfillment rates, and processing times**.

---

### 2 Inventory Optimization & Demand Forecasting

📌 **Concepts Used:**

- **ABC Analysis** → Categorizes inventory into fast-moving, slow-moving, and obsolete items.

- **Safety Stock Optimization** → Ensures buffer stock is set dynamically based on demand fluctuations.

- **Time-Series Forecasting** (ARIMA, Prophet, LSTM) → Predicts future demand trends to prevent stockouts.

📊 **Example Application:**

- **Stockout Reduction**: Using **machine learning** to improve demand forecasts and optimize reorder points.

- **Warehouse Slotting Optimization**: Placing frequently ordered SKUs near high-traffic picking zones.

## 3️⃣ Warehouse Process Optimization & Automation

📌 **Concepts Used:**

- **Lean Six Sigma** → Eliminates inefficiencies (waste in storage, picking, and packing).

- **Warehouse Layout Optimization** → Uses **clustering algorithms (K-Means, DBSCAN)** to group high-demand SKUs together.

- **Robotic Process Automation (RPA)** → Automates repetitive tasks like sorting and packing.

- **Digital Twin Technology** → Creates a real-time warehouse simulation to test layout and efficiency improvements.

📊 **Example Application:**

- **Reducing Picking Time**: Implementing **zone picking and robotic pickers** to optimize worker movement.

- **Optimizing Space Utilization**: Using AI-driven layout modeling to minimize **dead space and congestion**.

## 4️⃣ Order Fulfillment & Process Efficiency

📌 **Concepts Used:**

- **Order Batching Algorithms** → Groups similar orders together to reduce travel time.

- *Shortest Path Algorithms (Dijkstra's, A)** → Optimizes picking routes for workers.

- **Predictive Labor Scheduling** → Uses **regression models** to adjust labor shifts based on demand surges.

📊 **Example Application:**

- **Reducing Delivery Delays**: Implementing AI-powered **order routing and dynamic workforce allocation**.

- **Minimizing Packing Errors**: Using **computer vision and barcode scanning** for automated verification.

---

**5 Supply Chain & Logistics Optimization**

📌 **Concepts Used:**

- **Network Flow Optimization** → Ensures inventory is allocated efficiently across distribution centers.

- **IoT & RFID Tracking** → Real-time tracking of inventory movement and shipment progress.

- **Last-Mile Delivery Optimization** → AI-driven route planning to **reduce shipping delays and fuel costs**.

📊 **Example Application:**

- **Dynamic Inventory Allocation**: Redistributing stock across fulfillment centers to **minimize shipping delays**.

- **Reducing Lost Shipments**: Implementing **RFID-based tracking and anomaly detection**.

---

**6 Predictive & Prescriptive Analytics**

📌 **Concepts Used:**

- **Predictive Analytics (Regression, Decision Trees, XGBoost)** → Forecasts warehouse demand and labor needs.

- **Prescriptive Analytics (Optimization Models, Linear Programming)** → Suggests best warehouse configuration and inventory allocation.

📊 **Example Application:**

- **Identifying Bottlenecks**: Using **process mining** to track inefficient workflows.

- **Optimizing Cost Reduction**: Implementing AI-driven recommendations for **storage allocation and staffing levels**.

## Conclusion

After analysing Amazon's warehouse operations, the key findings highlight inefficiencies in **inventory management, order fulfilment, warehouse layout, labour productivity, and logistics**. These inefficiencies contribute to **stockouts, delivery delays, and rising operational costs**.

By leveraging **data-driven solutions**, such as **demand forecasting, AI-powered automation, warehouse layout optimization, and predictive labour scheduling**, Amazon can significantly **enhance efficiency, reduce costs, and improve customer satisfaction**.

---

🔍 **Key Takeaways & Solutions**

| Identified Problem | Root Cause | Data-Driven Solution | Expected Impact |
|---|---|---|---|
| **High Stockout Rate** | Poor demand forecasting, inaccurate inventory tracking | AI-powered demand forecasting, real-time inventory monitoring | ↓ 30-50% stockouts |
| **Slow Order Fulfilment** | Inefficient picking routes, high processing time | Optimized warehouse layout, robotic pickers, shortest-path algorithms | ↓ 40% processing time |
| **High Labor Costs & Idle Time** | Inefficient workforce scheduling | Predictive labor allocation & automation | ↓ 25% labor costs |
| **Underutilized Warehouse Space** | Poor slotting strategies, congestion in high-demand zones | AI-driven space optimization & heatmap analysis | ↑ 20% space utilization |
| **Delivery Delays** | Inefficient last-mile logistics, delayed | AI-driven routing, automated sorting & | ↓ 35% |

| Identified Problem | Root Cause | Data-Driven Solution | Expected Impact |
|---|---|---|---|
| | picking & packing | packing | delivery time |

🚀 **Business Impact & Next Steps**

☑ **Faster Order Fulfilment** → Customers receive products sooner, increasing satisfaction

☑ **Lower Operational Costs** → More efficient labor allocation and space utilization

☑ **Higher Inventory Accuracy** → Reduces losses from overstocking or stockouts

☑ **Optimized Warehouse Layout** → Minimizes unnecessary movement and boosts productivity

☑ **AI-Driven Automation** → Streamlines warehouse operations and increases throughput

By continuously monitoring **real-time dashboards**, improving **predictive models**, and scaling **automation**, Amazon can **sustain long-term efficiency and cost savings**.

**amazon**

---

## THANK YOU

## For Any QUERY CONTACT US on

✉ suhane79jaya@gmail.com

**in** **https://www.linkedin.com/in/jaya-suhane-776142275/**

Ⓟ **Peerlist.io/jayasuhane/**