

FMCG Warehouse: Optimising Amazon's Distribution Efficiency

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.

Problem Statement

Amazon is experiencing inefficiencies in its warehouse operations. These inefficiencies 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 optimise warehouse performance.

Why is it Important to Solve? Addressing these inefficiencies is crucial for improving customer satisfaction, reducing operational costs, and ensuring timely delivery of products. Efficient warehouse operations are critical for maintaining a competitive edge in the retail market.

Project scope

The project aims to identify and address inefficiencies in warehouse operations across various regions to resolve issues such as stockouts, delivery delays, and increased operational costs. By analysing data on warehouse capacity, refill requests, transport issues, and certifications, the goal is to develop metrics for evaluating performance and propose actionable improvements. Key tasks include understanding performance indicators, generating insights, and recommending solutions to enhance distribution efficiency. The findings will be summarised in a report for management to drive operational enhancements and maintain a competitive edge in the retail market.

Methodology

Data Sources: The data used in this project was provided by our team and sourced from Kaggle.

Data Wrangling: Data cleaning was performed in Python to handle missing values, normalise formats, and ensure consistency across all variables.

Data Analysis: The analysis was conducted using both Python and SQL Server to calculate average revenue by industry, assess industry representation, identify geographic trends, and examine correlations between emergency visits and readmission rates.

Data Visualization: Data was visualised using PowerBI and Python. Various charts and graphs were created to effectively communicate findings on revenue distribution, industry growth rates, geographic trends, and other key insights.

Goals and KPIs

Goals

1. **Improve Warehouse Efficiency**
 - Enhance operational workflows to minimise stockouts and delivery delays.
2. **Increase Customer Satisfaction**
 - Ensure timely deliveries and maintain high product availability.
3. **Optimise Inventory Management**
 - Implement better inventory tracking and refill processes to reduce waste.

KPIs

1. **Stockout Rate**
 - Target: Reduce stockout rate by 15% over the next quarter.
2. **On-Time Delivery Rate**
 - Achieve a 95% on-time delivery rate.
3. **Operational Cost per Unit**
 - Decrease operational costs per unit by 10% within six months.
4. **Warehouse Utilisation Rate**
 - Maintain an 80% warehouse capacity utilisation rate.

Technical Processes

- Tools used: Python, PowerBi, SQL Server
- Clauses used in SQL: SELECT, GROUP BY, FROM, ORDER BY, SQRT, TOP, Covariance. LIMIT, AVG, COUNT, FETCH.
- Python library: Numpy, Pandas. Used Matplotlib, Plotly and Seaborn for visualisation in Python
- PowerBi: Used for visualisation and making reports.

Business Concepts Used

Market Understanding: Analyse warehouse data to identify key trends and regional differences in warehouse operations.

Customer Demographic: Focus on zones with the highest retail shop presence, such as the North zone, to understand regional demand.

Customer Behaviour: Track refill requests and storage issues to identify patterns and areas needing operational improvements.

Customer Retention: Address stockout and delivery delays to enhance customer satisfaction and loyalty.

New Customer Acquisition: Optimise warehouse efficiency to reduce costs and improve service, attracting new customers to Amazon's FMCG products.

Recommended Analysis:

Q.1 What is the central Data Tendency capacity size of the warehouses?

- The average capacity size of the warehouses is as follows: large warehouses have an average count of 5375, mid-sized warehouses have an average count of 5306, and small warehouses have an average count of 2438. This data indicates that large and mid-sized warehouses are more common compared to small warehouses within the network.

Q.2 How many warehouses are located in urban areas versus rural areas?

- The data reveals a significant difference in the number of warehouses located in urban versus rural areas. Specifically, there are 11,994 warehouses situated in rural locations, indicating a substantial presence in these areas. In contrast, urban areas host 1,125 warehouses. This disparity underscores a strong preference for rural locations for warehouse operations, possibly due to factors such as space availability and cost considerations.

Q.3 What is the total number of retail shops served by each zone?

- The North zone serves the highest number of retail shops at 27.3 million, followed by South with 16.3 million, West with 20.7 million, and East with 1.1 million.
- This distribution highlights significant differences in the number of retail shops across zones, with the North zone having the most extensive retail network.

Q.4 Calculate the average number of workers per warehouse.

- The average number of workers per warehouse varies slightly across sizes, ranging from about 28.79 to 28.95 workers.
- These averages provide insight into the workforce distribution efficiency based on warehouse size categories.

Q.5 Determine the percentage of warehouses with electric supply.

- About 65.12% of warehouses have an electric supply, indicating that the majority of warehouses are equipped with this essential service.
- This figure highlights a strong presence of electric supply infrastructure across the warehouse network.

Q.6 What is the average distance of warehouses zone & regional zones from the central distribution hub?

- The North zone has the shortest average distance from the central distribution hub at 161.95 km.
- The East zone has the longest average distance at 166.44 km.
- Distances across zones are relatively close, indicating that warehouse locations are fairly evenly distributed around the central hub.
- This distribution reflects how central the hub is to the various regional zones and helps in understanding the logistics and transportation efficiency.

Q.7 How many warehouses have reported storage issues in the last 3 months zone & regional zones also showcase the percentage.

- Overall Impact: A total of 24,515 issues have been reported across all warehouses.
- Zone Distribution: The North zone reports the highest number of issues with 5,202, while the East zone reports the fewest with 220 issues.
- Regional Zone Distribution: Zone 6 has the highest number of issues with 4,218, whereas Zone 1 has the fewest with 980 issues.

Q.8 Identify the top 3 zones with the highest number of refill requests in the last 3 months.

- Zone 6 has the highest number of refill requests, with 24,061.
- Zone 5 follows with 13,190 refill requests.
- Zone 4 is third with 12,286 requests.

These figures highlight the zones with the most significant demand for refills, indicating where refill management and inventory planning may need the most attention.

Q.9 Calculate the average number of government checks per zone & regional zones in the last 3 months.

- **Highest Average Checks:** **East** has the highest average number of government checks across zones, at 20.33 checks.
- **Lowest Average Checks:** **West** has the lowest average at 16.99 checks.
- **Regional Zone Insights:** **Zone 4** shows the highest average number of government checks at 22.72, while **Zone 2** has the lowest average at 14.35.
- This data helps identify where the most and least government checks are processed, guiding resource allocation and policy adjustments.

Q.10 Determine the most common type of government certification among warehouses.

- The most common type of government certification among warehouses is **Certification C**, with a total of 2,801 warehouses holding this certification.
- **Certification B** follows as the second most common, with 2,586 warehouses.
- **Certification C** is the most prevalent type of government certification among warehouses, indicating that a significant portion of warehouses fall into this category.

Conclusion:

Based on the analysis of Amazon's FMCG warehouse data, it's clear that optimising warehouse operations is crucial. By addressing stockouts, improving delivery times, and refining inventory management, we can significantly enhance efficiency and customer satisfaction. Focusing on key zones and leveraging insights from refill requests and storage issues will be instrumental in driving these improvements. Overall, a strategic approach to managing warehouse capacity and ensuring robust electric supply infrastructure will help maintain Amazon's competitive edge in the market.

Project owner

Name: Ritika Pathak

Email id: ritupat707@gmail.com

Peerlist: https://peerlist.io/ritika_p

Kaggle: <https://www.kaggle.com/ritika01mp>

Discord Username: [r.k.0.7](#)