A logo for college computing

Description automatically generated

**Assessment**

|  |  |
| --- | --- |
| *Marluce Taciana Bora* |  |
| *Student Number: 2024141* |  |
| *Module Title: Strategic Thinking* |  |
| *Optimization of Warehouse Inventory Process Using Data Analytics* |  |
| *Assessment Due Date: 29/03/2024* |  |
| *Date of Submission: 29/03/2024* |  |

**Declaration**

By submitting this assessment, I confirm that I have read the CCT policy on academic misconduct and understand the implications of submitting work that is not my own or does not appropriately reference material taken from a third party or other source.

I declare it to be my own work and that all material from third parties has been appropriately referenced.

I further confirm that this work has not previously been submitted for assessment by myself or someone else in CCT College Dublin or any other higher education institution.

Abstract

This Capstone project aims to address the challenges faced by inventory processes in a distribution environment, with a focus on optimizing operational efficiency and customer satisfaction. Utilizing a dataset provided by the Mega Star Distribution Center, simulated within a large warehouse with over 60,000 storage positions, we intend to analyze and improve inventory management processes.

The dataset includes a detailed list of products, warehouse stocks, picking records, and receipt records. These data provide valuable insights into the current warehouse performance and can be explored to identify optimization opportunities.

The methodology adopted involves data collection, preprocessing, and exploratory analysis, followed by the application of advanced data modeling techniques such as ARIMA for demand forecasting and linear programming for inventory optimization. We aim to evaluate the effectiveness of the proposed solutions based on predefined metrics, such as average order picking time, inventory accuracy, and customer satisfaction.

At the end of the project, we expect to deliver a comprehensive report of the analyses conducted, along with practical recommendations for continuous improvements in warehousing and logistics processes. This project not only aims to address immediate warehouse efficiency issues but also contribute to the evolution of inventory management practices in large-scale distribution environments.

.

Contents

[1. Introduction 1](#_Toc162586335)

[2. Goals 1](#_Toc162586336)

[2.1 Inclusions: 1](#_Toc162586337)

[2.2 Exclusions: 2](#_Toc162586338)

[2.3 Project Boundaries: 2](#_Toc162586339)

[3. Scope: 2](#_Toc162586340)

[3.1 Methodology: 3](#_Toc162586341)

[4. Project Timeline: 4](#_Toc162586342)

[5. Data Source: 4](#_Toc162586343)

[6. Permission to Access Kaggle Dataset 5](#_Toc162586344)

[7. References: 5](#_Toc162586345)

# Introduction

In the environment of a distribution center, inventory processes face several challenges that affect operational efficiency, delivery delays, and consequently, customer satisfaction. This work aims to demonstrate how the use of the Capstone Project can apply data analysis techniques to identify and solve existing problems in warehouse processes, focusing on minimizing picking time, improving warehouse layout, increasing inventory accuracy, demand forecasting, and reducing delivery time by providing valuable insights, innovative solutions, practical improvements to optimize efficiency, reduce costs, and improve overall process performance.

# Goals

• Conduct a detailed analysis of warehouse processes and workflow.

Thoroughly evaluate all warehouse processes, from receiving to shipping, identifying inefficiencies and bottlenecks.

• Identify key issues impacting inventory process efficiency.

Analyze available data to pinpoint major obstacles affecting inventory process efficiency, such as picking time, inventory accuracy, and inadequate workflow.

• Propose solutions to optimize overall warehouse process efficiency.

Develop and implement strategies to address identified issues, aiming to improve operational efficiency, reduce wait times, and increase inventory accuracy.

• Utilize data analysis techniques for demand forecasting and inventory planning optimization.

Apply machine learning algorithms, such as ARIMA for demand forecasting and linear programming for stock optimization, to ensure more precise and efficient stock management.

• Provide practical insights for continuous improvements and advancement of warehousing and logistics practices.

Interpret data analysis results to offer practical recommendations and innovative solutions that can be implemented for ongoing improvements in warehousing and logistics processes.

These objectives are crucial to meeting the project's needs, which focus on the practical application of data analysis techniques to improve inventory management in a warehouse environment. Each goal has been formulated to address specific aspects of the project, from initial analysis to solution implementation and the provision of recommendations for continuous improvement.

# Inclusions:

* Comprehensive analysis of warehouse processes and workflow.
* Identification and resolution of key issues affecting inventory process efficiency.
* Proposal and implementation of solutions to optimize overall warehouse process efficiency.
* Utilization of data analysis techniques for demand forecasting and inventory planning optimization.
* Provision of practical insights for continuous improvements in warehousing and logistics practices.

# Exclusions:

* Implementation of physical changes to the warehouse infrastructure.
* Direct involvement in supply chain management beyond the scope of warehouse processes.
* Development of software or hardware solutions outside the realm of data analysis.

# Project Boundaries:

The project will focus solely on the analysis and optimization of warehouse processes.

No alterations to the dataset structure will be made; analysis will be conducted using the provided data.

The project does not extend to operational changes or decisions outside the scope of inventory management.

# Scope:

The heart of the project lies in the application of advanced data analysis techniques. This includes data cleaning and preparation to modeling and interpretation of results. Through machine learning algorithms such as ARIMA for demand forecasting and linear programming for inventory optimization, precise and effective models will be developed to improve warehouse management.

• Data Collection:

Utilization of the dataset provided by the Distribution Center, containing information on products, warehouse stocks, picking records, and receipt records.

Exploration and understanding of available data, identifying its structure, quality, and relevance to project objectives.

• Data Preprocessing:

Cleaning and preparing data for analysis, including handling missing values, standardizing formats, and integrating relevant datasets.

Integration of datasets: merging different data sources to gain a comprehensive view of the inventory process.

• Exploratory Data Analysis (EDA):

Exploring data to identify trends, patterns, and anomalies in inventory, picking, and receipt records.

Using graphs, histograms, and descriptive statistics to understand data distribution and variability.

• Data Modeling:

Application of data modeling techniques for demand forecasting, inventory planning optimization, and pattern identification in warehousing processes, using time series models such as ARIMA to predict future product demand.

Inventory planning optimization using optimization algorithms, such as linear programming, to determine the optimal amount of stock to be maintained for each product.

• Evaluation and Interpretation of Results:

Evaluating the effectiveness of proposed solutions based on metrics such as average time required to fulfill an order, accuracy percentage in stock quantity, and customer satisfaction, assessing service quality based on feedback and satisfaction metrics.

Interpreting results to identify insights and practical recommendations to optimize inventory processes.

# Methodology:

• Demand Forecasting:

Time Series Models: Utilization of algorithms like ARIMA (Autoregressive Integrated Moving Average) or exponential smoothing models to forecast future product demand based on historical data.

"Time series models such as ARIMA (Autoregressive Integrated Moving Average) are powerful tools for forecasting future product demand based on historical data. ARIMA is a statistical model that takes into account autocorrelation in data and trend over time. By adjusting model parameters to fit historical data, we can make accurate predictions about future demand, allowing companies to adequately prepare to meet market needs." Rob J Hyndman and George Athanasopoulos

• Inventory Forecasting and Replenishment:

Utilization of machine learning models such as linear regression or neural networks to predict stock levels of different products and optimize stock replenishment based on forecasted demand.

• Inventory Planning Optimization:

Linear Programming: Implementing linear programming algorithms to determine the optimal amount of stock to be maintained for each product, considering constraints such as available physical space, storage costs, and desired service levels.

As cited by Hamdy A. Taha in his book, "To optimize inventory planning, it is essential to use linear programming algorithms. These algorithms can determine the optimal amount of stock to be maintained for each product, considering various constraints such as available physical space in the warehouse, associated storage costs, and desired service levels. By modeling the inventory planning problem as a linear programming problem, companies can make more informed and effective decisions about how to allocate their inventory resources, ensuring a balance between product availability and associated maintenance costs." Hamdy A. Taha

• Pattern Identification in Warehousing Process:

Clustering: Grouping products with similar behaviors in terms of seasonal demand, physical characteristics, or movement speed in the warehouse. This can aid in efficient resource allocation and defining storage strategies.

• Process Performance Analysis:

Decision Trees: Building decision tree models to identify key factors affecting inventory process performance, such as picking time, stock accuracy, and customer satisfaction.

• Warehouse Layout Improvement:

Monte Carlo Simulation: Using Monte Carlo simulation techniques to evaluate different warehouse layouts and identify the one that optimizes workflow, minimizing distances traveled by pickers and reducing picking time. This technique allows managers to create virtual warehouse models and test various configurations, identifying the one that optimizes workflow and reduces picking time. Simulation provides valuable insights into how layout adjustments can positively impact logistical operations, enabling more informed and strategic decision-making to improve warehouse performance.

# Project Timeline:

Phase 1: Preparation and Data Collection (First Month)

Identify and access relevant data sources

Collect and organize necessary data for analysis

Phase 2: Exploratory Data Analysis (First and Second Month)

Perform initial data cleaning

Explore data and identify trends

Utilize visualization techniques and statistical analysis

Phase 3: Data Modeling (Third Month)

Apply data modeling techniques (ARIMA, linear programming)

Develop and fine-tune models

Phase 4: Evaluation and Interpretation of Results (Third Month)

Evaluate the effectiveness of proposed solutions

Interpret results and document conclusions:

By adhering to this comprehensive scope, delineating clear boundaries, and outlining planned methods and approaches, the Capstone Project aims to provide valuable insights and recommendations for optimizing warehouse efficiency through data analysis.

# Data Source:

The chosen dataset is available on the Kaggle website at the link https://www.kaggle.com/datasets/aikinomichi/mega-star-distribution-centre?resource=download, providing information from a fictional company, simulated within a warehouse of over 500 square meters with more than 60 thousand positions for storing various products. The dataset provides information such as:

Product list containing important information such as product name, description, SKU, batch, and dimensions, i.e., relevant data for effective inventory management within the warehouse.

Warehouse stock, providing a detailed inventory report of products, showing available positions for warehouse products and their respective quantities housed in each position. Inventory is one of the most important parts of the supply chain because it defines product entry demand, manages stock replenishment, and optimizes storage space.

Picking records document the last five days of highest demand for separation, providing information on picking locations, separated products, quantities separated, and information on who was responsible for the picking separation. Analyzing this data can optimize picking processes, improve picking team performance, optimize layout, reduce separation errors, and consequently improve overall customer satisfaction.

Receipt records track new stock receiving activities in the warehouse, received shipments, products, quantities, origin tracking, and the person responsible for the receiving task and product entry into the system/warehouse. Analysis of this data aims to provide insights into process efficiency and identify possible bottlenecks.

The Mega Star Distribution Center Dataset offers a rich and diverse collection of data, providing an immersive experience in inventory management in a large-scale distribution center. This dataset presents an opportunity to gain practical insights and refine skills in inventory management. Although its scenarios are simulated, the presented dataset aims to contribute to continuous improvement and advancement of warehousing and logistics practices.

# Permission to Access Kaggle Dataset

The dataset used in this project was obtained from Kaggle, a platform for sharing datasets and data science competitions. By accessing and using this dataset, permission and usage guidelines established by the dataset creators and the Kaggle platform were observed and followed.

Dataset Name: Mega Star Distribution Center Dataset

Creator or Provider: Leon Nguyen

Source: Kaggle (https://www.kaggle.com/)

Link to Dataset: https://www.kaggle.com/datasets/aikinomichi/mega-star-distribution-centre?resource=download

License or Terms of Use: Available under Database Contents License (DbCL) v1.0 — Open Data Commons: legal tools for open data

• Description of Terms of Use:

The terms of use associated with this dataset specify that it may only be used for educational and research purposes, and requires proper attribution to the dataset creators. Commercial use is not permitted without prior permission from the creators.

• Compliance Statement with Guidelines:

We declare that all permission and usage guidelines established by the dataset creators and the Kaggle platform were followed during access, analysis, and utilization of this dataset in this project.

# References:

Taha, H. A. (2007). Operations Research: An Introduction. Pearson Education.

Chopra, S., & Meindl, P. (2015). Supply Chain Management: Strategy, Planning, and Operation. Pearson Education.

Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: Principles and Practice. OTexts.

Provost, F., & Fawcett, T. (2013). Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking. O'Reilly Media.

Kaggle. (n.d.). Retrieved from https://www.kaggle.com/