|  |
| --- |
| BELGIUM CAMPUS |
| MLG382 PROJECT CYO |
|  |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  | | --- | --- | | Student | Student Number | | Karl Christiaan Schmutz | 577511 | | Noah Blaauw | 601195 | | Gordon Mullin | 600248 | | Eduard Jacobus Engelbrecht | 600237 |   14/04/2025 |

Contents

[Introduction 2](#_Toc196019944)

[Research Problem 3](#_Toc196019945)

[Dataset Details 4](#_Toc196019946)

[Dataset (Cleaned) 5](#_Toc196019947)

[Hypothesis Generation 6](#_Toc196019948)

[H0 (Null Hypothesis): 6](#_Toc196019949)

[H1 (Alternative Hypothesis): 6](#_Toc196019950)

[Correlation Data 7](#_Toc196019951)

[Features that increase profits: 7](#_Toc196019952)

[Features that decrease profits: 7](#_Toc196019953)

[ML Model(s) 8](#_Toc196019954)

[Web Application 9](#_Toc196019955)

[Development Process 10](#_Toc196019956)

[Deployment Steps 11](#_Toc196019957)

[Conclusion 12](#_Toc196019958)

[References 13](#_Toc196019959)

# Introduction

For small to medium businesses Inventory Management has become a hard requirement (Praveen et al., 2020), the importance of effective inventory management cannot be underestimated. It allows for operational efficiency and improving customer satisfaction whilst, traditional approaches frequently prove to be lacking in addressing the issues associated with fluctuating demand, variable supplier lead times, and the associated risks of overstocking or stockouts. The application of machine learning presents a comprehensive solution by harnessing historical sales data, analysing market trends, and evaluating supply chain variables. Using these predictions organizations can mitigate risks. According to (Praveen et al., 2020) high competition, labour unrest and changes in governmental laws can be addressed with predictions for products and services. This produces precise demand forecasts and optimizes inventory levels. Applying an inventory management methodology gives predictive insights and encourages better decision making (Chaudhary et al., 2023). Inventory management then changes from a reactive measure to a proactive strategy.

# Research Problem

Vertex PC Supply faces multiple internal inventory management challenges that significantly impact both operational efficiency and customer satisfaction. Despite collecting vast amounts of sales and inventory data, the company struggles to understand how to change this information into a proactive response to stop these issues from happening. Below are the Key issues discussed:

* Frequent Stockouts of High-Demand Items:
  + Incorrect demand forecasting often leads to empty shelves during peak demand period, resulting in lost sales and a decline in customer trust (Chaudhary et al., 2023).
* Excess Inventory and Elevated Carrying Costs:
  + Over purchasing of certain items causes surplus inventory, and a demand for capital incurring unnecessary storage expenses (Chaudhary et al., 2023).
* Low Profitability in Certain Regions or Categories:
  + According to (Sudirjo, 2023), global markets show that companies adapt their strategies to align with the unique consumer preferences, legal requirements and market conditions. This explains why a product will have different sales worldwide.
* Cost and Complexity:
  + Designing machine learning models can take a long time before it is effective and beneficial to the business. Costs are impacted not only by the time required for the model to be developed and deployed but also the expertise required to design it (Chaudhary et al., 2023).

The aim of this research is to design and implement a machine learning framework that predicts profitability based on key product, regional, and pricing attributes. By identifying what makes profit increase and decrease, Vertex PC Supply can optimize inventory decisions, regional marketing and distribution, reduce excess inventory and minimize stockouts.

# Dataset Details

The dataset utilized for this research contains various data from Vertex PC Supply, which provides essential variables that significantly impact inventory management decisions. Following are attributes within the dataset:

|  |  |
| --- | --- |
| Attribute | Description |
| RegionName | Region of operation |
| CountryName | Country of transaction |
| State | State within country |
| City | City of transaction |
| WarehouseName | Warehouse fulfilling order |
| CategoryName | Product category (e.g., Video Card) |
| ProductName | Specific product name |
| ProductStandardCost | Standard unit cost |
| Profit | Profit per unit |
| ProductListPrice | Product list price |
| CustomerCreditLimit | Customer’s credit limit |
| Status | Order status |
| OrderDate | Date of order |
| OrderItemQuantity | Quantity per order item |
| PerUnitPrice | Price paid per unit |
| TotalItemQuantity | Total items in order |

The dataset provided by [Kaggle](https://www.kaggle.com/datasets/hetulparmar/inventory-management-dataset) captures essential details across geographic locations, product specifications, financial metrics and order details. It includes attributes such as region, country, and warehouse to track the flow of inventory, alongside product categories and specific component names for precise analysis. Pricing elements like standard cost, list price, and per-unit price help in financial evaluation, whilst quantities and order dates enable demand forecasting and trend analysis. Additional fields such as customer credit limits and order status support customer management and operational efficiency. Combining these attributes provides a comprehensive view to effectively apply machine learning for inventory optimization and decision making.

## Dataset (Cleaned)

|  |  |
| --- | --- |
| Attribute | Description |
| WarehouseName | Warehouse fulfilling order |
| ProductName | Specific product name |
| ProductStandardCost | Standard unit cost |
| Profit | Profit per unit |
| ProductListPrice | Product list price |
| CustomerCreditLimit | Customer’s credit limit |
| Status | Order status |
| OrderItemQuantity | Quantity per order item |
| PerUnitPrice | Price paid per unit |
| TotalItemQuantity | Total items in order |
| Year | Derived feature from OrderDate |
| Month | Derived feature from OrderDate |
| Day | Derived feature from OrderDate |
| DayOfWeek | Derived feature from OrderDate |
| CostPriceRatio | Derived feature from ProductStandardCost/ProductListPrice |
| RegionName\_Australia | One-hot encoded region name |
| RegionName\_North America | One-hot encoded region name |
| RegionName\_South America | One-hot encoded region name |
| RegionName\_Canada | One-hot encoded region name |
| RegionName\_China | One-hot encoded region name |
| RegionName\_India | One-hot encoded region name |
| RegionName\_Mexico | One-hot encoded region name |
| CountryName\_United States of America | One-hot encoded country name |
| State\_California | One-hot encoded state |
| State\_Distrito Federal | One-hot encoded state |
| State\_Maharashtra | One-hot encoded state |
| State\_New Jersey | One-hot encoded state |
| State\_New South Wales | One-hot encoded state |
| State Ontario | One-hot encoded state |
| State\_Texas | One-hot encoded state |
| State\_Washington | One-hot encoded state |
| City\_Bombay | One-hot encoded city |
| City\_Mexico City | One-hot encoded city |
| City\_Seattle | One-hot encoded city |
| City\_South Brunswick | One-hot encoded city |
| City\_South San Francisco | One-hot encoded city |
| City\_Southlake | One-hot encoded city |
| City\_Sydney | One-hot encoded city |
| City\_Toronto | One-hot encoded city |
| CategoryName\_Mother Board | One-hot encoded product category |
| CategoryName\_RAM | One-hot encoded product category |
| CategoryName\_Storage | One-hot encoded product category |
| CategoryName\_Video Card | One-hot encoded product category |

# Hypothesis Generation

## H0 (Null Hypothesis):

There is no significant correlation between product pricing, product category, regional factors and overall profitability.

## H1 (Alternative Hypothesis):

Profitability is significantly affected by factors such as product list price, standard cost, regional factors. An example of this is that certain products and specific markets tend to have an increase in profits, where other products and regions would have a negative effect on profitability.

# Correlation Data

## Features that increase profits:

|  |  |  |
| --- | --- | --- |
| Feature | Correlation | Description |
| Product List Price | +0.85 | Selling higher-priced products brings significantly more profit. |
| Product Standard Cost | +0.77 | More expensive source products (high end tech) yield better profits when sold correctly. |
| USA as Country | +0.59 | U.S. market is a major profit centre. More sales should be targeted here. |
| Category: Video Card | +0.38 | These are among the most profitable items. Increase marketing and inventory for these items. |
| State: New Jersey | +0.34 | Customers in this state are consistently providing higher profits. |

## Features that decrease profits:

|  |  |  |
| --- | --- | --- |
| Feature | Correlation | Description |
| Category: Mother Board | -0.48 | These items generate poor profits. Investigate inventory and pricing |
| Region: Australia / City: Sydney / State: NSW | -0.41 | Sales in this region drag down profit. Investigate reasons (low demand, taxes, shipping costs) |
| Total Item Quantity | -0.40 | Selling large quantities lowers profits due to bulk discounts or shopping costs |

# ML Model(s)

# Web Application

# Development Process

# Deployment Steps

# Conclusion

# References

Chaudhary, V. et al. (2023) 'Exploring the Use of Machine Learning in Inventory Management for Increased Profitability', *New Zealand Herpetology*, 12(1), pp.658-66.

Praveen, K.B. et al. (2020) 'Inventory Management using Machine Learning', *International Journal of Engineering Research & Technology (IJERT)*, 9(06), pp.866-69.

Sudirjo, F. (2023) 'Marketing Strategy in Improving Product Competitiveness in the Global Market', *Journal of Contemporary Administration and Management (ADMAN)*, 1(2), pp.63-69.