www.idosr.org

Katushabe *et al* 2024

©IDOSR PUBLICATIONS

ISSN: 2550-794X

International Digital Organization for Scientific Research

IDOSRJSR1.1.5171.100

IDOSR JOURNAL OF SCIENTIFIC RESEARCH 9(1) 51-71, 2024.

https://doi.org/10.59298/IDOSRJSR/2024/1.1.5171.100

Enhancing Sales Demand Prediction for Supply Chain Management: A Dimensionality Reduction Approach at Mukwano Company Limited, Uganda

Katushabe Mercy[1], Muhaise Hussein[1], Akampurira Paul[1] and Val Hyginus Udoka Eze[2,*]

[1]**Department of Computing, Kampala International University, Uganda.**

[2]**Department of Electrical, Telecommunication and Computer Engineering, Kampala International University, Uganda.**

***Corresponding Author: Val Hyginus Udoka Eze, ezehyginusudoka@gmail.com, udoka.eze@kiu.ac.ug. Kampala International University, Western Campus, Ishaka, Uganda**

ABSTRACT

In the dynamic landscape of today's business environment, the accurate prediction of sales demand plays a pivotal role for companies striving to maintain competitiveness. This study focuses on refining demand sales prediction for supply chain management at Mukwano Industries Limited through the application of dimensionality reduction techniques. The research involved a systematic iterative process encompassing problem identification, data preparation, modeling, evaluation, validation, optimization, and documentation. To ensure confidentiality, secondary data from the Mukwano IT department underwent meticulous merging and anonymization. Four dimensionality reduction algorithms, namely PCA, SVD, MDS, and t-SNE, were employed and evaluated using RMSE metrics. The results reveal that MDS and t-SNE exhibited exceptional performance, achieving accuracies of 89% and 88.8%, respectively. PCA and SVD also demonstrated commendable performance with an accuracy of 82.4%. The study underscores the crucial role of dimensionality reduction in enhancing predictive accuracy and optimizing inventory management. Recommendations include the incorporation of feature selection and regularization techniques to address the challenges associated with the curse of dimensionality. In conclusion, this research contributes valuable insights into the effectiveness of diverse dimensionality reduction techniques for demand prediction and inventory management. Additionally, the study highlights the need to address the curse of dimensionality and suggests exploring further research avenues in other aspects of supply chain management. These recommendations are essential for guiding future research efforts in this evolving field.

Keywords: Demand sales prediction, Supply chain management, Feature selection, Inventory management, Predictive accuracy.

INTRODUCTION

In recent years, the utilization of Big Data Analytics (BDA) has emerged as a considerable asset for decision support systems, prompting Supply Chain (SC) managers to integrate these innovative techniques into their operations. The application of advanced BDA within the supply chain, known as Supply Chain Analytics (SCA), is characterized by three primary branches: Descriptive analytics, Predictive analytics, and Prescriptive analytics [1]. Supply chain management (SCM) revolves around the efficient flow of information, products, and services from origin points to end consumers through a network of interconnected organizations and activities [2]. While capacity, demand, and cost are recognized as standard variables in typical SCM challenges, uncertainties arise from shifts in client demand, supply chain dynamics, organizational risks, and lead times. Notably, demand uncertainty significantly influences SC performance, impacting production scheduling, inventory management, and transportation arrangements [3]. To maintain competitiveness and enhance profit margins, organizations are increasingly embracing precision marketing strategies. Consequently, forecasting models have become integral in precision marketing, enabling a comprehensive understanding and implementation of optimal inventory management strategies to meet customer demands and expectations [4][5]. In addressing the challenges posed by supply chain uncertainty, the strategic implementation of demand forecasting emerges as a pivotal solution. Within supply chain management (SCM), demand forecasting employs various statistical analytic methods, such as time-series analysis and regression analysis, as highlighted by [3][5]. The supply chain generates vast and high-dimensional data at multiple points, encompassing factors like items, supplier capacity, orders, shipments, and interactions with consumers and retailers. The continuous and rapid execution of numerous transactions across the supply chain network further amplifies this complexity [6]. Given the diverse range of suppliers, goods, and consumers involved, there has been a notable departure from conventional statistical demand forecasting methods. These methods typically rely on identifying statistically significant trends, characterized by mean and variance attributes in historical data. Instead, there is a growing preference for intelligent forecasting approaches that dynamically adapt by learning from past experiences. This shift is crucial for accurately anticipating the ever-shifting demands within supply chains [7]. To achieve this adaptability, big data analytics play a pivotal role by uncovering links between demand data across supply chain networks and deriving forecasting rules from these interactions. The implementation of such approaches involves the utilization of sophisticated machine-programmed algorithms, although these come at a computational cost [3]. Despite their complexity, these advanced analytics empower supply chain professionals to navigate the intricacies of demand forecasting in a more intelligent and responsive manner.

Theoretically, the Technology Acceptance Model (TAM) serves as a theoretical framework for assessing the likelihood of a system being embraced by potential users. An individual's inclination towards adopting a system hinges on its ability to enhance supply chain performance. Even if an employee initially lacks interest in a system, there's a high probability of adoption if the system facilitates the dissemination of emergency safety messages among staff members. Given the multitude of systems offering similar functionalities, users are likely to gravitate towards those that are more user-friendly and convenient, as opposed to ones that prove cumbersome [8].

Ease of use, a pivotal aspect, manifests through two essential mechanisms: instrumentality and self-efficacy. Simplifying system usage enhances users' sense of efficacy, fostering a belief that they have more control over their actions. The correlation between efficacy and ease of use directly influences attitudes, forming the intrinsic motivation for adoption. Aligning a system closely with user requirements during the design phase, ease of use ensures usability and aids in achieving specific goals such as efficiency, satisfaction, and effectiveness throughout the application development process. Recognized as a paramount factor, usability should not be deferred until the final product delivery; rather, it must be integrated into every stage of application development to meet customer expectations [9][10]. Conceptually, Big data analytics has evolved as a formidable strategy for enhancing predictive accuracy to better align with customer demands and optimize supply chain performance. It serves to assess supply chain efficiency, diminish response times, and fortify risk evaluation by refining inventory management practices [11]. Employing machine learning and data analytics algorithms enables the generation of precise demand forecasts, rooted in data, leading to the alignment of supply chain activities. This not only improves efficiency and customer satisfaction but also minimizes overall supply costs [12]. In the domain of supply chain management (SCM), there is a growing emphasis on the integration of big data analytics (BDA) for demand forecasting. Meta-research studies have delved into the applications of BDA in SCM, shedding light on the advantages, challenges, and gaps in demand forecasting within supply chains [3]. The convergence of big data, advanced analytics, and cloud services has streamlined the handling of vast datasets, opening new avenues for data-driven demand forecasting and planning [13]. Leveraging data mining algorithms and tools significantly enhances demand prediction accuracy, resulting in scalable forecasting models tailored for effective supply chain management [14]. A critical facet in the evolution of demand forecasting models is the exploration and development of an optimal model through data dimensionality reduction techniques. Dimensionality reduction aids in curtailing the number of features, extracting vital characteristics, and amalgamating similar features to amplify model performance [15]. For example, in the printed circuit board industry, the successful implementation of K-means clustering has proven effective in eliminating noise from input data and enhancing sales forecasting precision [16]. Researchers have underscored the utility of machine learning (ML) in demand forecasting, employing models like Regression Trees to predict sales, providing insights into the pivotal factors influencing sales and quantifying the impact of diverse actions [17]. Additionally, a prescriptive approach has been recommended, allowing managers to adjust performance improvement parameters based on model recommendations. This approach aids in identifying areas for enhancement and offers guidance on the extent of performance improvements [18][19]. The incorporation of ML and prescriptive analytics empowers organizations to leverage their expertise and knowledge acquired through trial-and-error experiences, facilitating informed decision-making in supply chain management [20]. Contextually, with cutting-edge technology that spans all industries, Mukwano Industries aims to provide high quality and reasonably priced Fast Moving Consumer Goods (FMCGs) such as packaged drinking water, toilet soaps, edible oils and fats, powder detergents, domestic and industrial plastics, and proteins for animal feeds. Through the entire process of supply chain, a lot of digital data is collected and stored in data warehouses. Moreover, there is an exponential increase in the data on a daily basis because most trade is done based on online transactions. This data is used among other things for marketing and analytical purposes towards on demand sales and keeping an optimal supply chain [21]. This shift to digital buyer journeys also comes with an increase in the sheer amount of information put in front of both the suppliers and buyers every moment they're online. Because of the upsurge in digital noise, the effect on the chain supply and management is overwhelming where it is difficult to glean insights from the vast Big data for on-demand sales prediction towards better stock management and an efficient supply chain [22]. Supply chain management (SCM) is concerned with the movement of products, services, and information from points of origin to consumers via a network of interconnected organizations and activities. It is believed that capacity, demand, and cost are well-known variables in typical SCM challenges. However, this is not the case in practice since lead times, organizational risks, and changes in consumer demand, supply, and transportation are all unknown [23]. Demand uncertainty in particular has a significant impact on SC performance with wide-ranging implications for scheduling production, planning inventories, and organizing transportation. In this regard, demand forecasting is a crucial strategy for tackling supply chain uncertainty [24]. Due to the diversity of suppliers, goods, and consumers, supply chain data is produced at many points in the chain for a variety of reasons in high volumes and at a high velocity, which is reflected in the numerous transactions that are continually processed throughout supply chain networks [25]. In light of these complexities, there has been a shift away from traditional (statistical) demand forecasting methods that rely on finding statistically significant trends (characterized by mean and variance attributes across historical data) in favor of intelligent forecasts that can intelligently evolve by learning from the past and adjusting to anticipate the constantly shifting demand in supply chains [26]. Because there is so much data accessible, analyzing it to provide timely insight is essential. For many businesses, predicting future demand is crucial because it influences operational choices. It is difficult for managers to optimize stock or for marketers to offer personalization and it is difficult to scale their efforts. Additionally, teams that want to reach decision-makers must compete with a lot of noise from the data, and marketing and sales teams want to measure their success. This is because there are many features in the data that present the curse of dimensionality and inefficient forecasting given the amount of data available in the databases [3]. The answer is to use dimensionality reduction techniques to minimize the amount of features in the dataset while retaining (or even enhancing) the effectiveness of any prediction model [27]. The most popular and efficient dimensionality reduction techniques, however, include Principal component analysis, Factor analysis, t-distributed Stochastic Neighbor embedding (t-SNE), and Forward Feature Selection. Other dimensionality reduction techniques include Missing value ratio, Low Variance Filter, and Random Forests. Another challenge is having the aforementioned tools and being able to select the best one for the job. To develop a more accurate predictive model for on-demand sales, the researcher plans to examine various data dimensionality methodologies, run simulations, and conduct tests. A possible competitive advantage in many industries is the ability to forecast future demand to enhance company analysis and decision-making. Managers of businesses may react swiftly to shifting market signals and subsequently modify their procurement and production strategies by using precise predictions. When early adoption is used in reaction to increasing demand or expenses are reduced in response to a fall, these options may result in higher revenues [28]. Traditionally, databases are manually analyzed to find patterns or extract knowledge. Using the traditional method, data is manually examined for patterns in an effort to extract knowledge. To assist the analyst in deriving similar results or information from the data, artificial intelligence may be used. The process of identifying patterns within a database is referred to as data mining, information retrieval, knowledge extraction, etc. Knowledge discovery refers to the overall process of learning from data, whereas data mining refers to the use of different intelligence algorithms to find patterns from the data. Data preparation, data selection, data cleansing, and data visualization are further stages [29]. Extraction and modification of features. Data transformations have a surprisingly big effect on the outcomes of data mining techniques. In this respect, the features' composition and/or transformation is a more important determinant of the quality of the outputs of data mining. The majority of the time, the ability to compose features depends on an understanding of the application, and the preparation of data is much improved when feature composition activities are approached from many disciplinary perspectives. However, general-purpose methods like principal component analysis (PCA) are frequently employed and have a high rate of success [30][31]. When one wants to preserve the original meaning of the characteristics and decide which of those features are crucial, feature selection is often favored over extraction/transformation. In addition, just the chosen characteristics need to be computed or gathered, unlike transformation-based approaches where all input data are still required in order to achieve the reduced dimension. The human analyst may choose a subset of the characteristics present in the initial data set depending on their understanding of the application domain and the mining effort's objectives. The feature selection process can be human or assisted by a few automated processes. In general, one of the three conceptual frameworks the filter model, the wrapper model, and the embedded methods are used to apply feature selection methods [32]. The inclusion of the learning algorithm in the assessing and choosing of features varies between these three fundamental groups. Without explicitly attempting to improve the effectiveness of any particular data-mining approach, the filter model selects characteristics as a preprocessing step. This is often accomplished by choosing a subset of attributes that maximizes the function of a (ad hoc) evaluation function using a search strategy. Due of the abundance of initial characteristics, doing a comprehensive search is typically infeasible [33][34][35]. At Mukwano Industries, our Supply Chain department is dedicated to maximizing overall value generation while efficiently fulfilling client orders and maintaining well-managed stocks through real-time supply management. A mastery of the supply chain offers a significant competitive advantage, enabling companies to optimize value creation [34]. However, relying on traditional demand prediction tools like Excel at Mukwano has proven limiting, lacking the analytical power of business intelligence. This limitation results in inaccurate inventory management and challenges in meeting customer needs, often leading to financial losses and a diminished competitive edge [35][36][37]. To address this issue, our data warehouses house extensive datasets containing product and store information crucial for demand prediction. Yet, the abundance of features used in prediction introduces the challenge of the curse of dimensionality, causing noise and demanding more time for processing and modeling. Consequently, optimizing stock for personalized marketing becomes challenging, hindering scalability and making it difficult for teams to reach decision-makers amid data noise [36][37][38][39]. Marketing and sales teams face obstacles in measuring success due to difficulties in accessing inconsistent and unreliable data, leading to misallocated resources, customer losses, and product expirations [38].

Recognizing the urgency of addressing the overwhelming data available and the critical need for timely insights, this research implements an integrated approach to dimensionality reduction and predictive analytics. Through a comparative analysis of four dimensionality reduction techniques—Factor Analysis, Principal Component Analysis, Forward Feature Selection, and t-Distributed Stochastic Neighbor Embedding—the study aims to establish an optimized model for on-demand sales prediction. This model seeks to enhance stock and sales management, ensuring better service quality at Mukwano Industries Limited.

Methodology

To find suitable models and obtain results from the data, this study used simulation and testing through iterative methods. As a result, the procedures of issue identification, data preparation and comprehension, data modeling, model assessment, model validation, and model optimization, as well as finalization and documentation, were used to carry out our technique.

Research Design

The research design followed a quantitative approach, aiming to develop predictive models for on-demand sales using dimensionality reduction techniques. The study adopted a cross-sectional design, utilizing secondary data from Mukwano Industries data warehouses with management consent. This quantitative research design focused on extracting valuable insights from on-demand sales data by developing predictive models through dimensionality reduction. It followed a cross-sectional approach using secondary data from Mukwano Industries, guided by the CRISP-DM framework.

Data collection

Secondary data from Mukwano Industries data warehouses with the management consent and then used for study in our research. Therefore, the collected data was imported into an advance statistical and data mining tool simulation and experimentation.

a) Duplicates removal

It is problematic to waste space and runtime with duplicate rows. Duplicate rows create incoherence, and the ML model fails to learn new information and therefore these were neutralized.

b) Data scaling or normalization

When using any data mining or machine learning systems, normalization is a valuable tool. The learning phase may be accelerated by normalizing the data features in tanning faces for back propagation neural network methods. The researcher used the following methodologies to normalize the prepared dataset; Min-max normalization, Decimal scaling normalization, and Z-score normalization.

c) Data transformation

New features created by data transformation, also known as changing features, where mathematical formulae derived from business models or pure mathematical formulae were used to integrate the raw input features. Linear, quadratic, polynomial, non-polynomial, rank, Box-Cox transformations were few to name among different existing transformation techniques and these were important since normalizations was not sufficient in research experiments and full automation to fit the data and optimize the resulting model. In some circumstances, combining the data embedded in several features may be advantageous. Outlier detection and treatment: Statistical methods for detecting outliers include box plots, scatter plots, z-scores, and IQR (Interquartile Range) scores for other distributions; a percentile-based approach was used in which values were regarded as outliers.

d) Outlier treatment involved the following

Trimming: Outliers from the dataset were removed, and it's not recommended. Capping: In capping, if the values are greater or less than an established threshold, those are deemed outliers, and the outlier numbers in the dataset determine the capping number. Outlier Removal Clustering (ORC): Eliminates outliers in each loop, a modification of K-Means Clustering. ORC efficiently removed outliers from clusters. As the model precision changes depending on the dataset, the parameters were adjusted appropriately. ORC guarantees that the centroid computation was not biased by locations distant from the k-clusters.

Design dimensionality reduction models towards on demand sales prediction

Exploratory data analysis (EDA) and data reduction (DR)

In this process, the target or dependent column and independent features were obtained. The DR, EDA, and clustering techniques reduced runtime and space during the deep-learning modeling phase.

a) Identifying redundant features

Feature Redundancy lengthened the modeling time of ML algorithms and led to model overfitting. Feature redundancy arose from the possibility of derivation from another feature or set of features.

b) Covariance and correlation

In statistics, covariance refers to the amount that two features or factors change in tandem whose value lies in the range of: Positive covariance ($-\infty$, $+\infty$) indicates they move in the same direction. Negative covariance means that any features are eater than the mean, and others are less vice-versa. Zero covariance means features may be independent under a certain hypothesis.

c) Feature selection

The reasons for conducting FS may include: removal of unnecessary data, enhancing forecasting accuracy, data cost and model complexity reduction, improvement in training efficiencies such as reductions in space needs and computational costs

d) Feature extraction

Feature extraction accelerates the ML algorithm's execution, optimizes raw data quality, boosts the algorithm's efficiency, and simplifies the interpretation of the findings. These include; Principal component analysis (PCA), Factor analysis, Multidimensional scaling (MDS), Singular Value Decomposition (SVD).

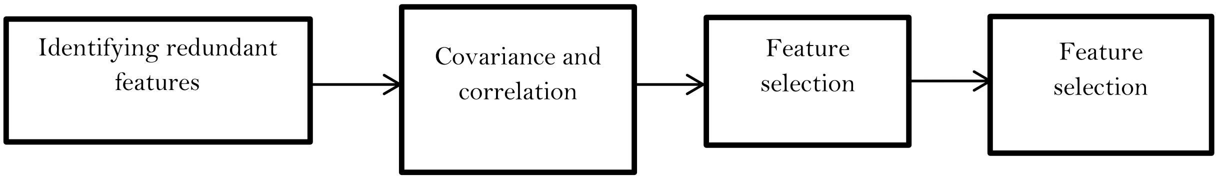

Figure 1: Process Model Block Diagram

Methods used to evaluate and validate the established model for on demand sales prediction for supply chain management at Mukwano industries limited are:

a) Model fit and train

The dataset was randomly split into train, validation, and test set for unbiased evaluation with new data to evaluate predictive performance with data different from training data. The best approach was to split a dataset by a date feature. The most recent samples were utilized for validation and testing. The primary concept was to choose a sample subset that accurately reflects the model data [40][41]. Two factors were used to determine the proportions of these three sets: the number of data samples and training models. Some models require a significant quantity of training data; therefore, the model was tuned for more extensive training sets in this scenario. Models with fewer hyper-parameters were easier to validate and tune, allowing a small validation set size. If the model contains considerably more significant hyper-parameters, having a large validation set was beneficial. There was no requirement of validation set if the model had no hyper-parameters or was challenging to adjust. When using k-fold CV, the train-test dataset splitting was repeated for k-times, with each new set being given a shot at becoming the hold-out set. Time-series data was not to be used with k-fold CV directly since they believe that there lies no connection between the rows and that they were all separate instances. For time-series data, instances' time-horizon prevents arbitrarily dividing them into clusters. Rather, data should be segmented, and the chronological sequence of instances maintained [42][43]. The term back testing was used in time-series forecasting to describe the technique of evaluating models using past data. In meteorology, this was regarded as 'hindcasting' rather than 'forecasting'.

b) Forecasting evaluation

The predicting abilities of the models were compared using five error metrics. Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) as show in equation (2.11) and (2.12), two well-known and established scale-dependent error metrics, made up the first two criterions. They were frequently used because they were simple to compute and comprehend. For retailing managers, they also have practical implications since they naturally give greater importance to fast-moving SKUs because they often generate more income than slower-moving products in a shop.

c) Model validation

The models' speed, resource utilization in terms of needed computing power, accuracy (taking error rate into account), precision, and recall were all taken into account. To assess the models' accuracy, ROC curves and confusion matrix techniques were mostly employed. Comparative analysis of the modeling phase's algorithmic performance was done during the model assessment phase. The main goal of this was to identify the algorithms' strengths and shortcomings in order to choose the best algorithm for demand sale forecast. A more suitable model would be inexpensive for a business, easily explicable (i.e., the algorithm's results would be clear and simple enough to aid in decision-making), and robust (i.e., able to be used in real-world scenarios without degrading performance). In order to assess the effectiveness of demand sales, the algorithms' accuracy, precision, recall, and F1measure on test data sets were taken into consideration. Parameter adjustment was done at this stage to finally optimize the model and prepare it for deployment, as well as validation trials to ensure that the selected method works better on untested instances.

Ethical Considerations

Throughout the course of this study, ethical considerations were meticulously upheld to ensure the integrity and confidentiality of all involved parties. The use of secondary data from Mukwano IT department was conducted in strict compliance with established privacy protocols. Stringent measures were implemented to anonymous and safeguard the data, thereby preserving the confidentiality of sensitive information.

Results and Data Interpretation

The researcher collected secondary data from Mukwano IT department in two parts which were merged in excel before being introduced to our analysis tool. The collected data was provided in categories and no real product names were provided and it was labeled with abstract variable names for confidentiality purposes before being sent out of the company databases. The data was imported into Jupiter notebook for exploration, visualizations, preparation analysis and modeling. The imported data is shown in Table 1.

Table 1: A snippet portion of the dataset

	QUANTITYORDERED	Price_in_shillings	Territor Key	SALES	ORDER_MONTH_ID	Order_year	Order_Demand
count	65535.000000	65535.000000	65535.000000	65535.000000	65535.000000	65535.000000	6.553500e+04
mean	35.084794	307268.187541	6.314092	3546.274548	7.097536	2015.816739	1.025558e+04
std	9.773471	71700.813523	2.960397	1838.284186	3.657087	0.737466	4.427651e+04
min	6.000000	98112.000000	1.000000	482.130000	1.000000	2015.000000	0.000000e+00
25%	27.000000	255901.500000	4.000000	2194.170000	4.000000	2015.000000	2.500000e+02
50%	35.000000	349670.000000	7.000000	3182.970000	8.000000	2016.000000	2.000000e+03
75%	43.000000	365000.000000	9.000000	4496.800000	11.000000	2016.000000	7.000000e+03
max	97.000000	365000.000000	10.000000	14082.800000	12.00000	2017.000000	4.000000e+06

The Table 1 shows a description of the dataset showing the count for each variable which is 65535 examples, the means and standard deviations and percentiles. To explore the different datatype inscribed in our dataset, we performed tests such as the data type and data count. The out of the test result is depicted in the snippet shown in Plates I(a)-I(b).

```
datta.dtypes
Out[5]:
Product_Category      object
Warehouse From        object
Stock_Date            object
QUANTITYORDERED       int64
Price_in_shillings    float64
 Total_Amount         object
CATEGORY              object
TerritoryKey          int64
SALES                 float64
ORDER_MONTH_ID        int64
Order_year            int64
Order_Demand          int64
STATUS                object
dtype: object
```

(a)

```
##datta.count()- Counting the number of records for each field
Out[6]:
Product_Category      65535
Warehouse From        65535
Stock_Date            65535
QUANTITYORDERED       65535
Price_in_shillings    65535
 Total_Amount         65535
CATEGORY              65535
TerritoryKey          65535
SALES                 65535
ORDER_MONTH_ID        65535
Order_year            65535
Order_Demand          65535
STATUS                65535
dtype: int64
```

(b)

Plate I: Output Result of Data type and Data Count Tests

The dataset shows that most of the data was stored as objects and hence descriptive analysis of the data can only be done after the data has been converted to discrete or continuous numerical data. Also as shown in Plate I, the total records under each of the features of our dataset showed that our data was ok with all records.

RESULTS OF THE STUDY TO ESTABLISH REQUIREMENTS FOR DESIGNING A PREDICTIVE MODEL FOR DIMENSIONALITY REDUCED DATASETS

Null values in dataset were checked to find out if some data cells were empty and null values would hinder modeling hence had to be delt with if they existed. The function #is null was used when checking Null values. This function takes a scalar or array-like object and indicates whether values are missing. There are no Null values in our dataset. The output result of checking the null, NaN and duplicate values is as shown in Plate II (a-c).

```
Product_Category      0
Warehouse From        0
Stock_Date            0
QUANTITYORDERED       0
Price_in_shillings    0
 Total_Amount         0
CATEGORY              0
TerritoryKey          0
SALES                 0
ORDER_MONTH_ID        0
Order_year            0
Order_Demand          0
STATUS                0
dtype: int64
```

(a)

```
# We need to check NaN values in dataset
datta.isna().sum()
Product_Category      0
Warehouse From        0
Stock_Date            0
QUANTITYORDERED       0
Price_in_shillings    0
 Total_Amount         0
CATEGORY              0
TerritoryKey          0
SALES                 0
ORDER_MONTH_ID        0
Order_year            0
Order_Demand          0
STATUS                0
dtype: int64
```

(b)

```
# To check duplicated values
print(datta.duplicated().value_counts())
False    65535
dtype: int64
```

(c)

Plate II: Output Result of checking the null, NaN and duplicate values

The Plate II shows that there were no NaN values in our dataset and was therefore good for further analysis and exploration and would be good for modeling. It also shows that there is no duplicate in our dataset and was good further explorations and analysis. With a view to visualizing the features, the general data plot was obtained as depicted in Figure 2. The Figure 2 shows the overall general plot of our data features. Figure 2 shows a good distribution of the data and shows that all the features' examples are visualized.

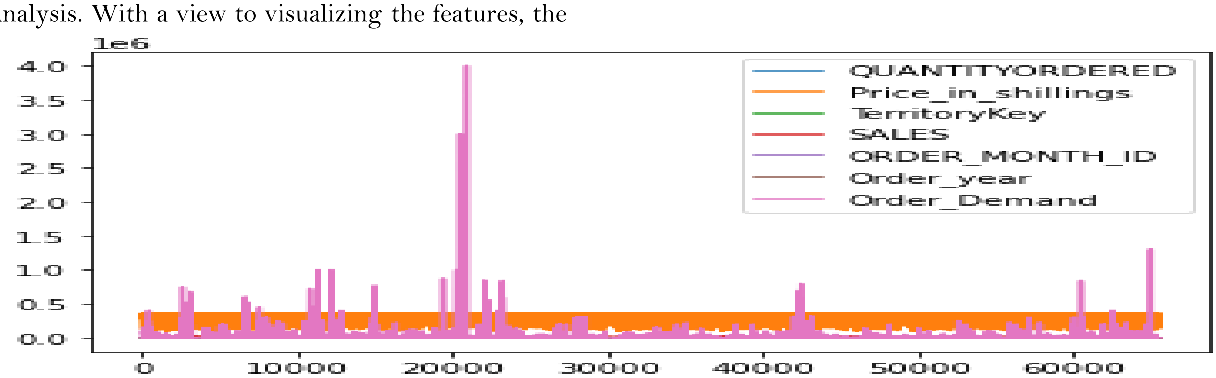

Figure 2: General Data Plot

The sales distribution plot was done. The order-month distribution analysis shows that data is normally distributed around the mean with most orders having been done in the month of November and the lowest orders received were in the months of June and July as shown in the Figure 3.

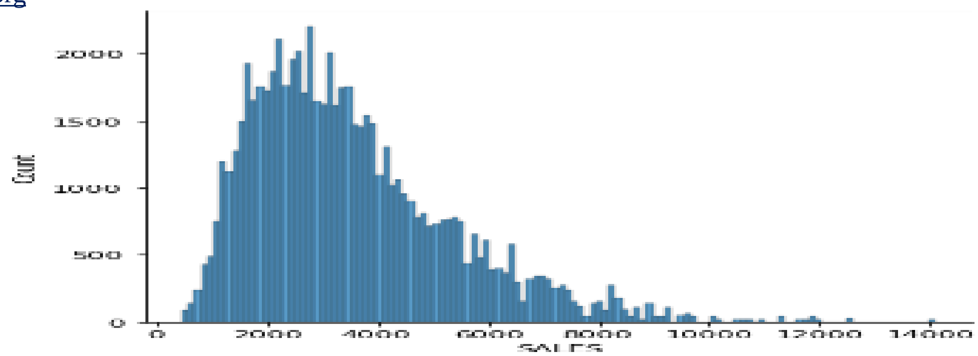

Figure 3: Sales Distribution Plot

The Figure 3 shows a well distribution of the Sales and shows a normal distribution with a skewness to the right of the mean. To know the Order Months Distribution, there is exploring of the different order processed from different warehouses as seen in the Figure 4.

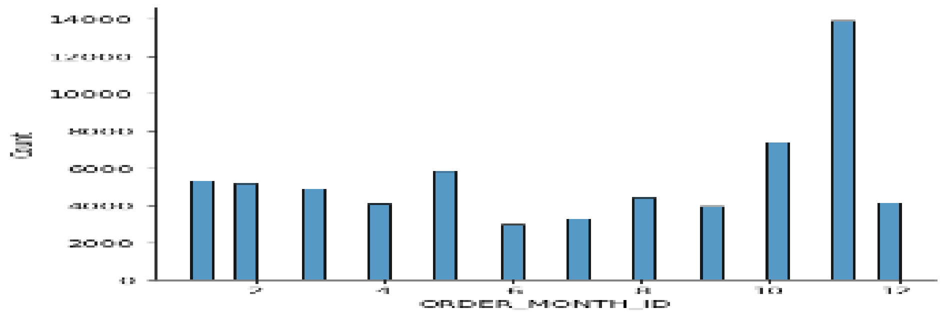

Figure 4: Order Months Distribution plot

In order to visually explore the status of the order received over a period of three years using a count plot as shown in Figure 5, a status order analysis was performed.

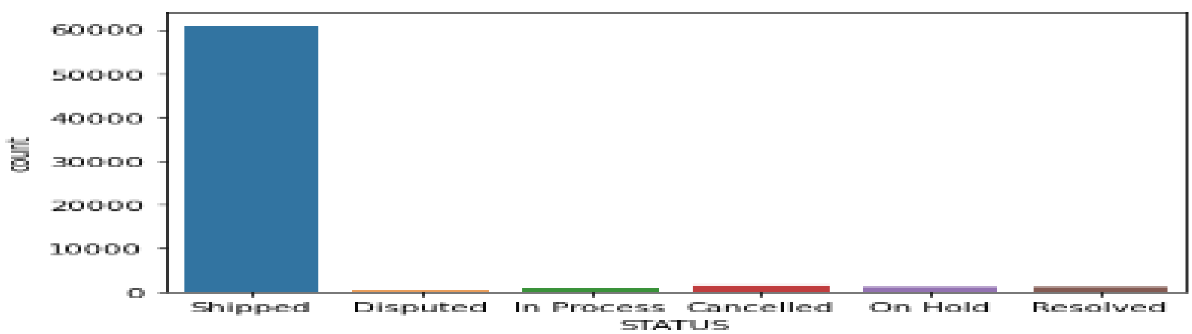

Figure 5: Count Plot for status of orders

The Figure 5 shows that of all the received orders, more than 60000 were successfully shipped to the clients and others fell in the categories of disputed, in process, cancelled, on hold and resolved. To know the warehouse inventory progress, there is exploring of the different order processed from different warehouses as seen in the Figure 6

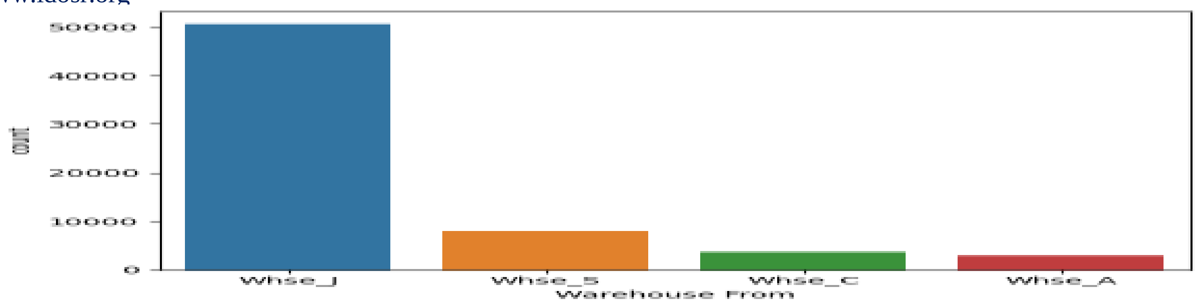

Figure 6: Number of samples according to Warehouse

The different labeled warehouses represent real ware houses and warehouse J was dominantly getting mos t orders, 50746, as compared to other warehouses w here Whse_S had 8081 orders, Whse_C had 3744 an d Whse_A had 2964. To know about the different re gions of order origin, we looked at the territory data as seen in the Figure 7.

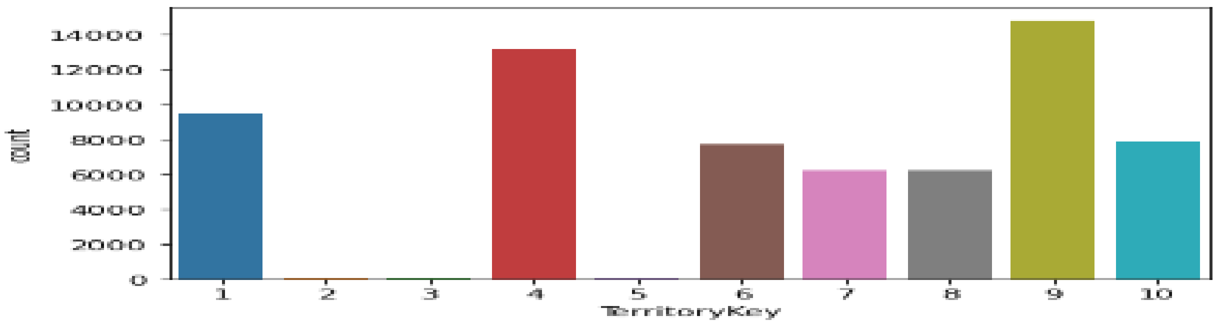

Figure 7: Number of samples according to Territory Key

The count plot in Figure 7 shows that territory 9 makes the most orders followed by territory four and one respectively. The territories with least orders include territories two, three and five. To know the different categories of products and how they sale, the histogram was used as shown in the Figure 8.

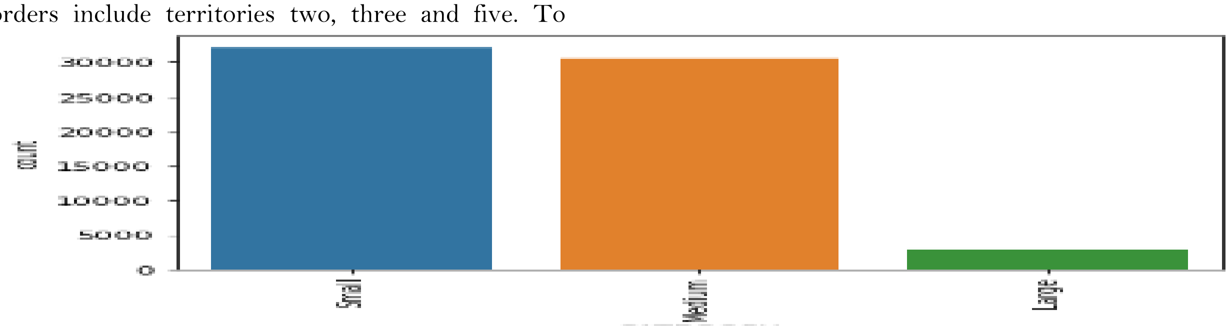

Figure 8: Product Category

The Figure 8 shows that product category small sales are the most, followed by medium size products and finally large category products sales are the least. Check of the distribution of the independent variables using Q-Q probability plots. Q–Q plot (quantile-quantile plot) is a probability plot, a graphical method for comparing two probability distributions by plotting their quantiles against each other;

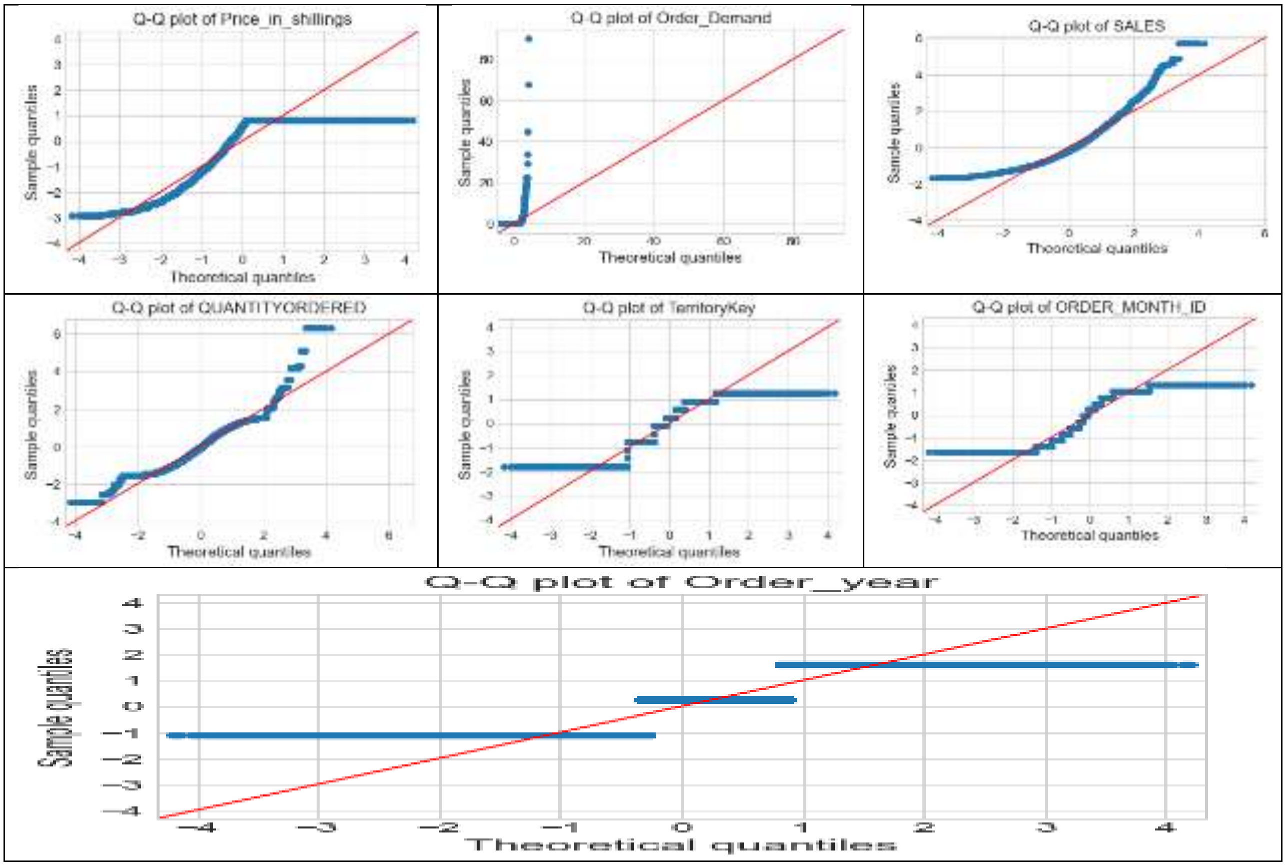

Figure 9: Sample quantities of theoretical quantiles

The Figure 9 shows the different distribution of quantiles shown for the features of the dataset and they mostly show linear grow effect of the features except for the order-demand which exhibited exponential growth rate.

Results of the study to design and develop the dimensionality reduction model for prediction of on demand sales

Data features and dummy categories were first labeled for the product categories as shown in table 2.

Table 2: Product Categories Encoded

	Product_Code	Warehouse From	QUANTITYORDERED	Price_in_shillings	CATEGORY	TerritoryKey	SALES	ORDER_MONTH_ID	Order_year	Ord
0	Product_0993	Whse_J	30	349305.0	Small	1	2871.00	2	2015	
1	Product_0979	Whse_J	34	296927.5	Small	9	2765.90	5	2015	
2	Product_0979	Whse_J	41	345801.0	Medium	1	3884.34	7	2015	
3	Product_0979	Whse_J	45	303899.0	Medium	4	3746.70	8	2015	
4	Product_0979	Whse_J	49	365000.0	Medium	9	5205.27	10	2015	

The dataset with dummy variables where categorical features were encoded with dummy variables as shown in the Table 3. This was done so as to ease the modeling process with quantitative variables instead.

Table 3: Dummy variables for categorical features

)RDERED	Price_in_shillings	TerritoryKey	SALES	ORDER_MONTH_ID	Order_year	Order_Demand	Product_Category	Warehouse From	CATEGORY	STATUS
30	349305.0	1	2871.00	2	2015	100	22	2	2	5
34	296927.5	9	2765.90	5	2015	500	22	2	2	5
41	345801.0	1	3884.34	7	2015	500	22	2	1	5
45	303899.0	4	3746.70	8	2015	500	22	2	1	5
49	365000.0	9	5205.27	10	2015	500	22	2	1	5

The methodology of Min-max normalization was applied on our dataset to rescale the entire feature so that they can have a same magnitude within our data set and each feature was scaled to range between 0 and 1 inclusive, as shown below.

Table 4. Normalized data

	QUANTITYORDERED	Price_in_shillings	TerritoryKey	SALES	ORDER_MONTH_ID	Order_Demand	Warehouse From	CATEGORY
0	0.263736	0.9570	0.000000	0.175644	0.090909	0.000025	0.666667	1.0
1	0.307692	0.8135	0.888889	0.167916	0.363636	0.000125	0.666667	1.0
2	0.384615	0.9474	0.000000	0.250150	0.545455	0.000125	0.666667	0.5
3	0.428571	0.8326	0.333333	0.240030	0.636364	0.000125	0.666667	0.5
4	0.472527	1.0000	0.888889	0.347273	0.818182	0.000125	0.666667	0.5

The Table 4 shows the scaled feature dataset which was then ready for further modeling and analysis.

Analysis was embarked to determine which independent variables are useful to predict a target (dependent variable (Sales)). Two methods were applied including, Pearson Correlation and Select K Best as shown in the table 5 and plate III

Table 5: Analysis to determine independent variables and Plate III Analysis to determine independent variables

	QUANTITYORDERED	Price_in_shillings	TerritoryKey	ORDER_MONTH_ID	Order_Demand	Warehouse From	CATEGORY	STATUS	Product_Category
QUANTITYORDERED	1.000000	-0.000212	0.005615	-0.004849	0.002929	0.004154	-0.023437	0.004444	0.004444
Price_in_shillings	-0.000212	1.000000	-0.006331	0.001215	0.013817	-0.063902	0.035503	0.003122	0.003122
TerritoryKey	0.005615	-0.006331	1.000000	0.006205	0.008997	0.006451	-0.018534	-0.002935	-0.002935
ORDER_MONTH_ID	-0.004849	0.001215	0.006205	1.000000	-0.001016	0.000092	0.000008	0.025498	0.025498
Order_Demand	0.002929	0.013817	0.008997	-0.001016	1.000000	0.068633	-0.015256	0.000023	0.000023
Warehouse From	0.004154	-0.063902	0.006451	0.000092	0.068633	1.000000	-0.070972	-0.000303	-0.000303
CATEGORY	-0.023437	0.035503	-0.018534	0.000008	-0.015256	-0.070972	1.000000	-0.001009	-0.001009
STATUS	0.004444	0.003122	-0.002935	0.025498	0.000023	-0.000303	-0.001009	1.000000	1.000000
Product_Category	0.004444	0.003122	-0.002935	0.025498	0.000023	-0.000303	-0.001009	1.000000	1.000000
SALES	0.018417	0.012543	0.002308	0.012178	0.015623	0.036857	-0.058047	-0.002117	-0.002117

```
datta.corr().unstack().sort_values().drop_duplicates()
CATEGORY            Warehouse From      -0.070972
Price_in_shillings  Warehouse From      -0.063902
SALES               CATEGORY            -0.058047
QUANTITYORDERED     CATEGORY            -0.023437
TerritoryKey        CATEGORY            -0.018534
Order_Demand        CATEGORY            -0.015256
TerritoryKey        Price_in_shillings  -0.006331
QUANTITYORDERED     ORDER_MONTH_ID      -0.004849
STATUS              TerritoryKey        -0.002935
SALES               STATUS              -0.002117
ORDER_MONTH_ID      Order_Demand        -0.001016
Product_Category    CATEGORY            -0.001009
STATUS              Warehouse From      -0.000303
Price_in_shillings  QUANTITYORDERED      0.000212
ORDER_MONTH_ID      CATEGORY             0.000008
Product_Category    Order_Demand         0.000023
ORDER_MONTH_ID      Warehouse From       0.000092
Price_in_shillings  ORDER_MONTH_ID       0.001215
SALES               TerritoryKey         0.002308
Order_Demand        QUANTITYORDERED      0.002929
Price_in_shillings  STATUS               0.003122
QUANTITYORDERED     Warehouse From       0.004154
```

The analysis showed that Price, warehouse, category territory key and month determine much about the demand and number of sales. Using Select K best method the variables in Figure 10 are basic features or variables for estimating sales. To model the data will require a lot of time with all the columns present so the columns were removed in a way that do not affect the output variable from the feature selection process. For example, status was removed from predictive features as show in the Table 6.

Table 6: Select KBest method for feature importance and selection

	QUANTITYORDERED	TerritoryKey	ORDER_MONTH_ID	CATEGORY	STATUS
0	0.263736	0.000000	0.090909	1.0	1.0
1	0.307692	0.888889	0.363636	1.0	1.0
2	0.384615	0.000000	0.545455	0.5	1.0
3	0.428571	0.333333	0.636364	0.5	1.0
4	0.472527	0.888889	0.818182	0.5	1.0

R^2 is computed without centering (uncentered) since the model does not contain a constant in the Plate IV

Table 7: Plate IV Identifying Predictor Significance and Plate V: Feature Significance

OLS Regression Results

Dep. Variable:	SALES	R-squared (uncentered):	0.770
Model:	OLS	Adj. R-squared (uncentered):	0.770
Method:	Least Squares	F-statistic:	5.470e+04
Date:	Thu, 05 Jan 2023	Prob (F-statistic):	0.00
Time:	12:36:53	Log-Likelihood:	-5.8836e+05
No. Observations:	65535	AIC:	1.177e+06
Df Residuals:	65531	BIC:	1.177e+06
Df Model:	4		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
QUANTITYORDERED	38.6754	0.599	64.593	0.000	37.502	39.849
TerritoryKey	68.8577	2.358	29.206	0.000	64.237	73.479
ORDER_MONTH_ID	57.3940	1.927	29.790	0.000	53.618	61.170
Warehouse From	633.1286	10.251	61.764	0.000	613.037	653.220

Omnibus:	10190.893	Durbin-Watson:	1.117
Prob(Omnibus):	0.000	Jarque-Bera (JB):	17713.474
Skew:	1.020	Prob(JB):	0.00
Kurtosis:	4.525	Cond. No.	51.7

The Table 7, shows Feature significance toward predicting target variable of sales. Select the most important features of Total quantity ordered, the Territory or region, the order month and warehouse which are the most important features. Intention to determine their effect or significance on the sales and results showed that they are indeed independent from each other and have individual significance. The total amount order for has the most significant effect. The highest error rate was given by the warehouse from with a high bias and the lowest error was given by total quantities ordered for. Standard Errors assume that the covariance matrix of the errors is correctly specified

Results of the study to evaluate and validate the established model for on demand sales prediction for supply chain management at Mukwano industries limited.

The complexity and accuracy of the machine learning model both drop when the number of data variables is decreased. However, having fewer features makes it simpler to explore, interpret, and analyze data. It also reduces the computing cost of machine learning methods.

a) Principle Component Analysis (PCA)

The goal of PCA, to put it simply, is to minimize the number of variables in a data collection while retaining as much data as feasible. The Principal Component Analysis (PCA) class implementation of the scikit-learn package is available for usage as a dimensionality reduction data transform. The number of desired dimensions in the transform's output may be configured using the "n components" option.

Running PCA.
Explained variance: 0.9944
Individual variance contributions:
0.907732701885
0.0254234063215
0.0110337289521
0.0070268753562
0.00628341600314
0.00611974772924
0.00535864726681
0.00453274019813
0.00362930731976
0.00294262288265

0.00239246012144
0.00177464330232
0.00166653490752
0.00151185737228
0.00148285378183
0.00129501399992
0.00120521481347
0.0010727631062
0.00103952906349
0.000841552767779

Better than 90% of the data is explained by a single principal component. Just a shade less than 99% of variance is explained by 15 components, which means that this dataset can be safely reduced to ~15 features. The matrix vector shows that a negative coefficient and represents a strong negative association or a strong negative correlation with the variables included in the PCA. The variables that have a positive influence on this component will have a negative influence on the overall component score, and vice versa such as price and number of sales.

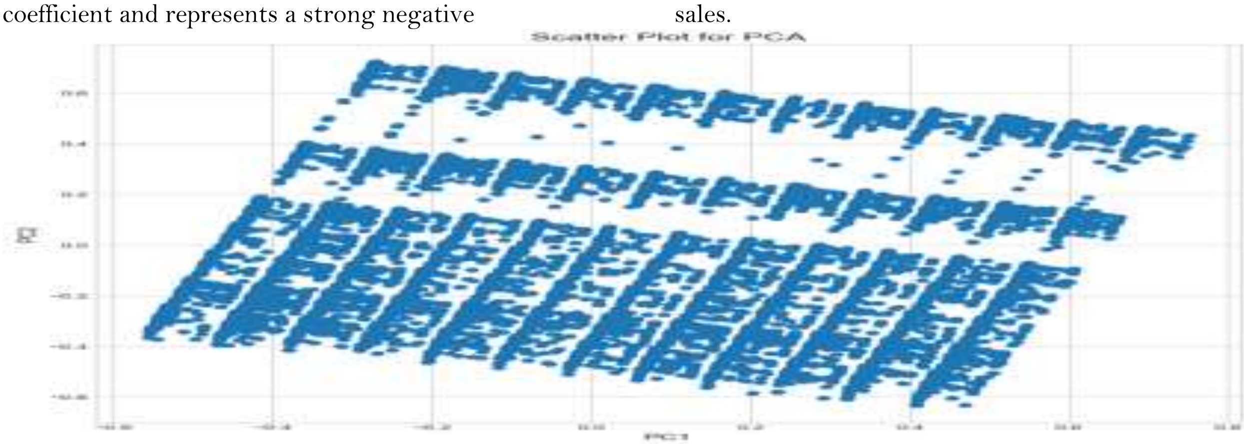

Figure 11: A Scatter plot for PCA

b) Singular Value Decomposition (SVD)

Singular Value Decomposition, or SVD, is one of the most popular techniques for dimensionality reduction for sparse data (data with many zero values). The scikit-learn library provides the Truncated SVD class implementation of Singular Value Decomposition that can be used as a dimensionality reduction data transform. The "*n_components*" argument can be set to configure the number of desired dimensions in the output of the transform

Accuracy: 0.824 (0.176)

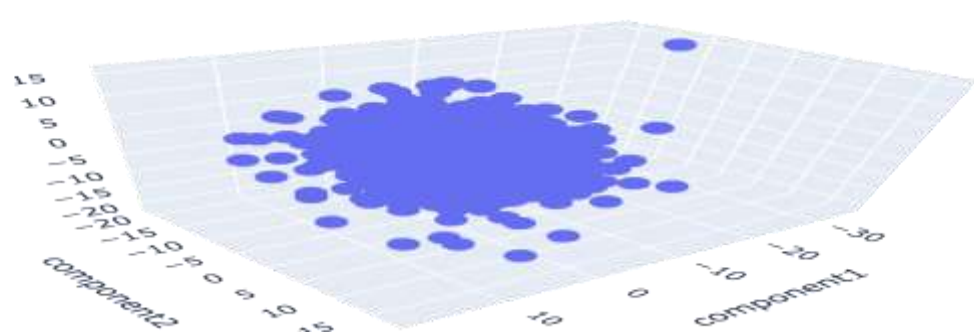

Figure 12: 3D plot for SVD

c) Multi-Dimensional Scaling (MDS)

MDS is similar to PCA. MDS is a Linear Dimensionality Reduction technique as well. The eigen vectors are found for the Dissimilarity matrix as opposed to the covariance matrix in PCA. MDS algorithm: Compute dissimilarity matrix Find eigen vectors and eigen values for the dissimilarity matrix. The eigen vector with high eigen value explains the information well that means the variance of the data is high Project the data points to the Eigen vector. Table 8 shows the dissimilarity Matrix of the Eigens.

Table 8: Eigen Vectors and Eigen Values dissimilarity Matrix

	Components1	Component2
0	-0.914207	0.510220
1	-0.912315	0.305411
2	-0.979451	0.064428
3	-0.979222	-0.050014
4	-0.980162	-0.137778

Component1 has a coefficient of -0.914207. The negative sign indicates an inverse relationship with the original variables.

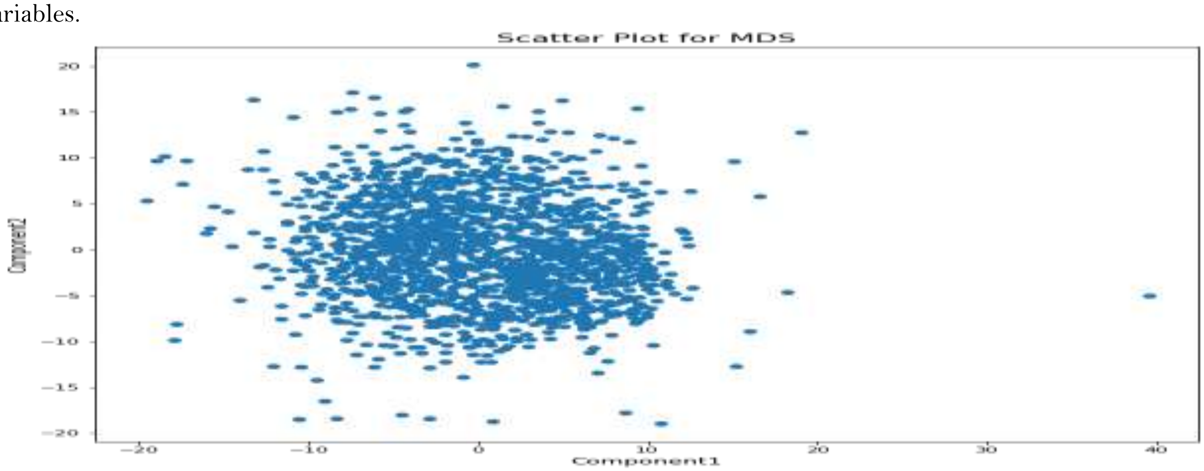

Figure 13: A Scatter plot for MDS

The MDS was able to reduce the dimensionality of our dataset and 91.4 of data explained by the first component and the first two components 96.5 of the variance explained.

d) T-distributed Stochastic Neighbor Embedding (TSNE)

A method called TSNE locates non-linear correlations in the data. It is a tool for data visualization. While PCA is concerned with maintaining large pairwise distances to optimize variance, T-SNE is focused with preserving only tiny pairwise distances or local similarities. In the beginning, t-SNE builds a probability distribution across pairs of high-dimensional objects in a way that gives comparable items a larger likelihood while giving different points a very low probability. Second, t-SNE minimizes the Kullback-Leibler divergence (KL divergence) between the two distributions with regard to the positions of the points in the map by defining a comparable probability distribution over the points in the low-dimensional map. Although the original approach bases its similarity metric on the Euclidean distance between objects, this may be modified as necessary. KL-Divergence, which is a measure of how two distributions differ, is also known as relative entropy. t-SNE could potentially lead to better data separation/visualization, because unlike PCA it preserves the local structure of data points. The problem with sklearn implementation of t-SNE is its lack of memory optimization. The "*n_components*" argument can be set to configure the number of desired dimensions in the output of the transform.

Accuracy: 0.888 (0.112)
Explained variance: 0.6231
Individual variance contributions:
0.074084070101
0.0618868203631
0.0559539511796
0.0427607611418

0.0359086048879
0.0345707250307
0.0314658500583
0.0285582078293
0.0255109028989
0.0252506880295
0.0245827886253
0.0241400322075
0.0238642769192
0.0232114066023
0.021035114187
0.0209135028964
0.020064977449
0.0189478800073
0.0162507118409
0.014169058539

t-SNE only managed to produce less desired variance explained of 63% as compared to other algorithms.

Table 9: t-SNE Accuracy

s/n	Component 1	Component 2
0	-68.874184	-12.161551
1	1.009264	0.037844
2	-29.502213	-41.437454
3	-6.174396	-29.133535
4	45.548256	-10.435360

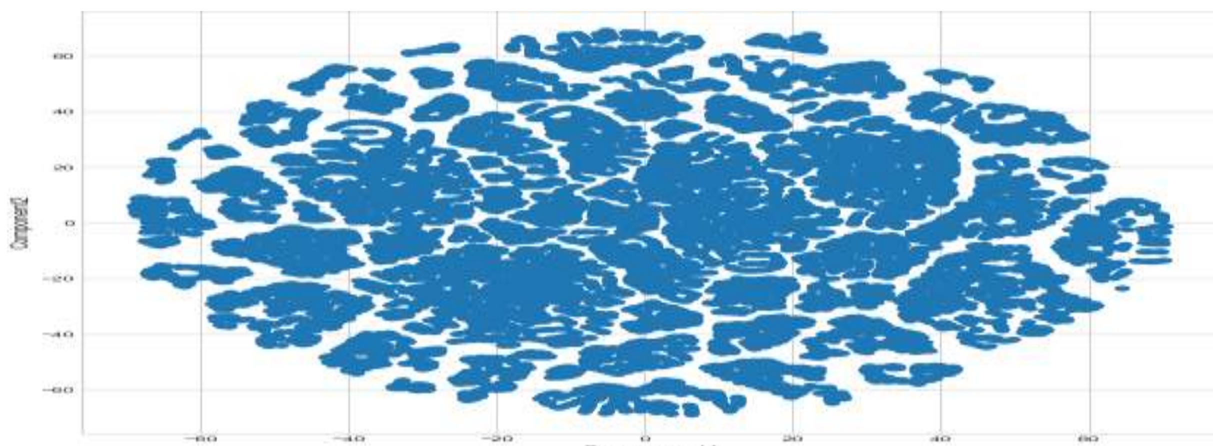

Figure 14: Scatter plot for t-SNE

The scatter plot shows a compact view of the variance explained where most is by the first two components. 80.9% of the variance is explained by the first two components. The results show that some components have an inverse relationship like cost and some show positive relationship like warehouse and category key.

***Comparative Analysis of PCA, t SNE, SVD and MDS**

A comparative analysis of the developed models shows that our best model was MDS with the least error of 0.108, followed by t-SNE with and error 0.112. The worse performer was SVD and PCA for our regression task but the performance was good enough since they managed to achieve 82.4% variance explainable. The comparative performance analysis of the dimensionality reduction algorithms indicates that MDS (Multi-Dimensional Scaling) achieved the best results among the models developed. It obtained the lowest error rate of 0.108, suggesting that it was able to preserve the original data structure more effectively compared to the other models. This indicates that the reduced-dimensional representation obtained by MDS closely resembles the original data, which is desirable in many applications. Following MDS, t-SNE (t-Distributed Stochastic Neighbor Embedding) performed relatively well with an error rate of 0.112. Although it had a slightly higher error compared to MDS, it still managed to capture the underlying patterns in the data effectively. t-SNE is known for its ability to preserve local structures and highlight clusters in high-dimensional data. Therefore, its performance in this analysis suggests that it was successful in capturing the important relationships among the data points. On the other hand, SVD (Singular Value Decomposition) and PCA (Principal Component Analysis) were identified as the worst performers for the regression task. However, their performance was still deemed acceptable as they achieved 82.4% variance explainable. This indicates that they were able to retain a significant portion of the data's variability in the reduced-dimensional representation. SVD and PCA are widely used dimensionality reduction techniques that focus on capturing the most important linear relationships in the data. While they might not have performed as well as MDS or t-SNE in this particular task, their ability to explain a high percentage of the variance suggests that they still provided valuable insights into the data. It's important to note that the choice of dimensionality reduction algorithm depends on the specific requirements of the task at hand. MDS and t-SNE are often favored for visualizing data and preserving local structures, while SVD and PCA are more commonly used for feature extraction and linear relationships. The results of this comparative analysis provide insights into the strengths and weaknesses of each algorithm, allowing researchers and practitioners to make informed decisions based on their specific needs and objectives.

Table 10: Comparative Performance Analysis of the Models

Model	*Accuracy*	*Error*
t-SNE	0.888	0.112
PCA	0.824	0.176
SVD	0.824	0.176
MDS	0.892	0.108

CONCLUSION

In conclusion, to prioritize dimensionality reduction techniques, feature selection, and regularization methods. Our research at Mukwano Industries Limited demonstrated the effectiveness of employing five dimensionality reduction algorithms, with MDS and t-SNE yielding exceptional accuracies of 89% and 88.8%, respectively. Furthermore, addressing the curse of dimensionality through feature selection and regularization techniques proved successful in mitigating challenges related to model size, complexity, and optimization times. The importance of aligned decision-making across the supply chain hierarchy cannot be overstated. Our research emphasized the need for tailored demand forecasts at different levels of aggregation within the supply chain. By ensuring decision-makers at various tiers have access to accurate and customized forecasts, organizations can enhance planning and resource allocation, ultimately optimizing supply chain management. Finally, the findings of this research underscores the critical role of demand predictions at multiple levels of aggregation for informed supply chain decision-making. When implementing machine learning approaches for inventory management and supply chain optimization, prioritizing dimensionality reduction, feature selection, and regularization techniques is paramount to achieving accurate and efficient results.

RECOMMENDATION

a) This study primarily focuses on developing a demand predictive model using dimensionality reduction techniques. It is imperative for the company to ensure the reliability, user-friendliness, and continuous availability of these models. Additionally, administrators should prioritize training instructors to effectively utilize the models, ensuring that the information/resources incorporated are both relevant and accurate.

b) The study utilized various models, including PCA, SVD, LDA, and t-SNE, all of which exhibited high performance accuracy rates. Notably, PCA and LDA demonstrated superior performance. To enhance the accuracy of t-SNE, it is suggested to optimize it further by reducing variance. However, to validate and build upon these findings, it is recommended to conduct additional comparative studies using diverse datasets and alternative dimensionality reduction algorithms.

c) The study identifies the curse of dimensionality as a challenge impacting model size, complexity, and optimization times. To address this issue, exploration of techniques such as feature selection and regularization is suggested. Machine learning methodologies and dimensionality reduction techniques are emphasized for their pivotal role in inventory management and supply chain optimization.

d) The study recommends the integration of domain-specific knowledge and expertise to enhance the accuracy and applicability of predictive models.

REFERENCES

1. Nguyen, T., Li, Z. H. O. U., Spiegler, V., Ieromonachou, P., & Lin, Y. (2018). Big data analytics in supply chain management: A state-of-the-art literature review. *Computers & Operations Research*, *98*, 254-264.
2. Fernando, Y., Tseng, M. L., Nur, G. M., Ikhsan, R. B., & Lim, M. K. (2023). Practising circular economy performance in Malaysia: managing supply chain disruption and technological innovation capability under industry 4.0. *International Journal of Logistics Research and Applications*, *26*(12), 1704-1727.
3. Seyedan, M., & Mafakheri, F. (2020). Predictive big data analytics for supply chain demand forecasting: methods, applications, and research opportunities. *Journal of Big Data*, *7*(1), 1-22.
4. Li, G., Zhao, Z., Zhang, K. N., Wang, Q., Zeng, F., Zhang, Y. & Jiang, T. (2021). Chinese Glioma Genome Atlas (CGGA): a comprehensive resource with functional genomic data from Chinese glioma patients. *Genomics, proteomics & bioinformatics*, *19*(1), 1-12.
5. Yazid, K., Silaji, T., Rahim, A., Eze, C. E., & Eze, V. H. U. (2023). The Effect of Financial Management on the Learning Ability of Students in Government-Aided Primary Schools in Ibanda Municipality Uganda. *International Journal of Humanities, Management and Social Science*, *6*(2), 109–118. https://doi.org/10.36079/lamintang.ij-humass-0602.600
6. Ying, L. (2020). An integrative framework of supply chain flexibility. *International Journal of Productivity and Performance Management*, *69*(6), 1321-1342.
7. Aktas, E. (2022). Big Data Applications in Supply Chain Management. In *The Palgrave Handbook of Supply Chain Management* (pp. 1-25). Cham: Springer International Publishing.
8. Hankun, H., Yafang, L., Xuemei, H., & Jing, F. (2016, June). A comparative study of China and US users' acceptance of online payment. In *2016 13th International Conference on Service Systems and Service Management (ICSSSM)* (pp. 1-6). IEEE.
9. Eric, D. F., Al Fadley, A. A., & Al Enezi, E. G. (2022). Exploring the attitudes of instructors toward Microsoft Teams using the Technology Acceptance Model. *International Education Studies*, *15*(1), 123-135.
10. Enyi, V. S., Eze, V. H. U., Ugwu, F. C., & Ogbonna, C. C. (2021). Path Loss Model Predictions for Different Gsm Networks in the University of Nigeria , Nsukka Campus Environment for Estimation of Propagation Loss. *International Journal of Advanced Research in Computer and Communication Engineering*, *10*(8), 108–115. https://doi.org/10.17148/IJARCCE.2021.10816
11. Wasike, C. L. (2020). *Big Data Analytics and Supply Chain Performance of Network Facilities Providers in Kenya* (Doctoral dissertation, University of Nairobi).
12. Kersten, W., Brylowski, M., Schroeder, M., Lodemann, S., & (2021, December). Machine learning in supply chain management: A scoping review. In *Hamburg International Conference of Logistics (HICL) 2021* (pp. 377-406).
13. Yang, C., Huang, Q., Li, Z., Liu, K., & Hu, F. (2017). Big Data and cloud computing: innovation opportunities and

challenges. *International Journal of Digital Earth*, *10*(1), 13-53.
14. Denis, R. R., & Blumenfeld, D. E. (2014). Product complexity and supply chain design. *International Journal of Production Research*, *52*(7), 1956-1969.
15. Mohsen, Baha, Big Data Application in Supply Chain Management: Scopus Based Literature Review. Available at SSRN: https://ssrn.com/abstract=3613716 or http://dx.doi.org/10.2139/ssrn.3613716
16. Chukwudi, O. F., Eze, V. H. U., & Ugwu, C. N (2023). A Review of Cross-Platform Document File Reader Using Speech Synthesis. *International Journal of Artificial Intelligence*, *10*(2), 104–111.
17. Venishetty, S. V. (2019). Machine learning approach for forecasting the sales of truck components, Blekinge Institute of Technology, Faculty of Computing, Department of Computer Science, DV2572 Master´s Thesis in Computer Science, 1-49.
18. Juan Pablo Usuga Cadavid, Samir Lamouri, Bernard Grabot. Trends in Machine Learning Applied to Demand & Sales Forecasting: A Review. International Conference on Information Systems, Logistics and Supply Chain, Jul 2018, Lyon, France. ⟨hal-01881362⟩
19. Eze, C. E., Eze, V. H. U., & Ugwu, J. (2024). Unveiling the Dynamics of E-Banking : A Comprehensive Analysis of Strategic Choices, Service Quality and Customer Satisfaction. *NEWPORT INTERNATIONAL JOURNAL OF CURRENT RESEARCH IN HUMANITIES AND SOCIAL SCIENCES*, *4*(1), 1–7.
20. Fildes, R. Trapero, J. R., & Kourentzes, N. (2012). Impact of information exchange on supplier forecasting performance. *Omega*, *40*(6), 738-747.
21. Stacy, M., Chicoine, N., Griffin, J., Ergun, O. (2023, May). Agent-Based Modeling of Human Decision-makers Under Uncertain Information During Supply Chain Shortages. In *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems* (pp. 1886-1894).
22. Seyedan, M., & Mafakheri, F. (2020). Predictive big data analytics for supply chain demand forecasting: methods, applications, and research opportunities. *Journal of Big Data*, *7*(1), 1-22.
23. Boulaksil, Y. (2016). Safety stock placement in supply chains with demand forecast updates. *Operations Research Perspectives*, *3*, 27-31.
24. Blackburn, R., Lurz, K., Priese, B., Göb, R., & Darkow, I. L. (2015). A predictive analytics approach for demand forecasting in the process industry. *International Transactions in Operational Research*, *22*(3), 407-428.
25. Mustafee, N., Katsaliaki, K., & Taylor, S. J. (2021). Distributed Approaches to Supply Chain Simulation: A Review. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, *31*(4), 1-31.
26. Michna.S, & Mafakheri, F. (2020). Predictive big data analytics for supply chain demand forecasting: methods, applications, and research opportunities. *Journal of Big Data*, *7*(1), 1-22.
27. Pulkit, V., Yadav, N., & Garg, P. (2018). American sign language fingerspelling using hybrid discrete wavelet transform-gabor filter and convolutional neural network. *Journal of Engineering Science and Technology*, *13*(9), 2655-2669.
28. Ivanov, N., Jovanović, R., Fichert, F., Strauss, A., Starita, S., Babić, O., & Pavlović, G. (2019). Coordinated capacity and demand management in a redesigned Air Traffic Management value-chain. *Journal of Air Transport Management*, *75*, 139-152.
29. Swarnalatha, P., & Sevugan, P. (Eds.). (2018). *Big data analytics for satellite image processing and remote sensing*. IGI Global.
30. Ogunleye, J. O. (2021). The Concept of Data Mining. In *Data Mining-Concepts and Applications*. IntechOpen.
31. Eze, V. H. U., Eze, M. C., Chidiebere, C. S., Ibokette, B. O., Ani, M., & Anike, U. P. (2016). Review of the Effects of Standard Deviation on Time and Frequency Response of Gaussian Filter. *International Journal of Scientific & Engineering Research*, *7*(9), 747–751.
32. Yinghui.S, Xu, H., Jiang, L., & Liu, Y. (2020). Few-shot modulation classification method based on feature dimension reduction and pseudo-label training. *IEEE Access*, *8*, 140411-140425.

33. Arif Wani, M., Kantardzic, M., Sayed-Mouchaweh, M. (2020). Trends in Deep Learning Applications. In: Wani, M., Kantardzic, M., Sayed-Mouchaweh, M. (eds) Deep Learning Applications. Advances in Intelligent Systems and Computing, vol 1098. Springer, Singapore. https://doi.org/10.1007/978-981-15-1816-4_.
34. Dash, S., Shakyawar, S. K., Sharma, M., & Kaushik, S. (2019). Big data in healthcare: management, analysis and future prospects. *Journal of big data*, *6*(1), 1-25.
35. Mukwano Industries reaps benefits of Cloud transformation, https://www.independent.co.ug/mukwano-industries-reaps-benefits-of-cloud-transformation.
36. V. C. Madala, S. Chandrasekaran and J. Bunk, "CNNs Avoid the Curse of Dimensionality by Learning on Patches," in IEEE Open Journal of Signal Processing, vol. 4, pp. 233-241, 2023, doi: 10.1109/OJSP.2023.3270082.
37. Eze, V. H. U., Eze, M. C., Enerst, E., & Eze, C. E. (2023). Design and Development of Effective Multi-Level Cache Memory Model. *International Journal of Recent Technology and Applied Science*, *5*(2), 54–64. https://doi.org/10.36079/lamintang.ijortas-0502.515
38. Eze, V. H. U., Ugwu, C. N., & Ugwuanyi, I. C. (2023). A Study of Cyber Security Threats, Challenges in Different Fields and its Prospective Solutions : A Review. *INOSR Journal of Scientific Research*, *9*(1), 13–24.
39. Ogenyi, F. C., Eze, V. H. U., & Ugwu, C. N. (2023). Navigating Challenges and Maximizing Benefits in the Integration of Information and Communication Technology in African Primary Schools. *International Journal of Humanities, Management and Social Science*, *6*(2), 101–108.https://doi.org/10.36079/lamintang.ij-humass-0602.599
40. Eze, V. H. U., Eze, M. C., Ogbonna, C. C., Valentine, S., Ugwu, S. A., & Eze, C. E. (2022). Review of the Implications of Uploading Unverified Dataset in A Data Banking Site (Case Study of Kaggle). *IDOSR Journal of Applied Science*, 7(1), 29–40.
41. Sreenivasulu Potluri & S. Swarnalatha (2023) An optimised resource allocation scheme based on NOMA-SWIPT in cooperative network, International Journal of Electronics, DOI: 10.1080/00207217.2023.2248567
42. Eze, V. H. U., Eze, C. E., Ugwu, C. N., Ogenyi, F. C., Ugwu, O. P., Obeagu, E. I., Alum, E. U., Okon, M. B., Ugwu, J. N., Obeagu, G. U., Mbabazi, A., Aleke, J. U., & Twesiime, M. (2023). Maximizing Journal Article Impact : Strategies for Enhanced Visibility in Today ' s Academic Sphere. *INOSR APPLIED SCIENCES*, *11*(1), 1–12.
43. Agarwal, R., Melnick, L., Frosst, N., Zhang, X., Lengerich, B., Caruana, R., & Hinton, G. E. (2021). Neural additive models: Interpretable machine learning with neural nets. *Advances in neural information processing systems*, *34(2021)*, 4699-4711.

CITE AS: Katushabe Mercy, Muhaise Hussein, Akampurira Paul and Val Hyginus Udoka Eze (2024). Enhancing Sales Demand Prediction for Supply Chain Management: A Dimensionality Reduction Approach at Mukwano Company Limited, Uganda. IDOSR JOURNAL OF SCIENTIFIC RESEARCH 9(1) 51-71. https://doi.org/10.59298/IDOSRJSR/2024/1.1.5171.100