VAAL UNIVERSITY OF TECHNOLOGY

VAAL UNIVERSITY
OF TECHNOLOGY

Inspiring thought. Shaping talent.

FACULTY:	Applied and Computer Sciences
DEPARTMENT:	Information Technology
SUBJECT:	Business Analysis 3.2
SUBJECT CODE:	AIBUY3A
LECTURER:	Mr. M Matsela
MODERATOR:	
YEAR:	2025

ASSESSMENT NAME:	B.A 3.2 Group Assignment
SUBMISSION DATE:	13/10/2025

STUDENT NUMBER	INITIALS AND SURNAME	SIGNATURE
223433861	US DUBE	
223371661	A MAUBANE	
222070560	P KHATHALI	
22307982	T RAMARA	

Table of Contents

LOGISTICS: WAREHOUSE MANAGEMENT SYSTEM

Background

In the industrial sector, warehouse operations play a crucial role in maintaining the efficiency of supply chains. By utilizing machine learning algorithms to forecast future customer demands, logistics companies can optimize their inventory, streamline supply chain operations, and enhance their overall responsiveness to customer needs. This solution will enable warehouses to reduce stockouts, overstocking, and costs, leading to improved customer satisfaction and increased competitiveness in the market. This aligns perfectly with the industrial theme, as it optimizes warehouse efficiency, reduces waste, and supports smarter, data-driven operations through the use of artificial intelligence.

Problem Definition

The challenges faced by Logistics companies demonstrate how difficult it can be to keep supply chains running smoothly. Many logistics and warehouse companies struggle with **inefficient demand forecasting** due to reliance on outdated manual methods. These incorrect forecasts cause either **overstocking**, leading to wasted storage space and higher costs, or **understocking**, resulting in delayed deliveries and unsatisfied customers.

Manual forecasting methods, relying on historical data and statistical models, are sometimes inaccurate and fail to account for constant changes/trends and seasonal fluctuations. This problem is particularly pressing in the industrial sector, where demand patterns can be complex.

Benefits to Local Municipality/Community

By introducing an AI warehouse management system, local municipalities and industrial companies can improve their operational efficiency, reduce costs, and enhance customer satisfaction. This, in return, can benefit the local community by ensuring the timely delivery of essential goods, supporting local businesses, and contributing to the overall economic growth and development.

Main Objectives

The main objective of our AI solution is to develop a demand forecasting model that can accurately predict future customer demands for logistics companies in the industrial sector. By leveraging machine learning algorithms and historical data, our solution aims to optimize inventory levels, reduce stockouts and overstocking, and improve supply chain operations.

Here are our key objectives:

- **Optimize Warehouse Operations:** To use data and AI-powered analytics to predict future customer demand and streamline all warehouse workflows. This will increase overall productivity and help the company save on costs related to labour and energy.
- **Analysing Logistics Data:** Use machine learning to analyse large datasets and provide insights into logistics operations, such as identifying trends, patterns, and areas for improvement.
- **Data Collection:** Gather past sales records, product movement logs, and external factors such as holidays and market trends.
- **Data Pre-processing:** Clean and structure the data for analysis, handling missing or inconsistent entries.
- **Model Training:** Use supervised learning techniques to train AI models on historical demand patterns.
- **Prediction:** Generate demand forecasts for various products and categories.

- **User-friendly:** Ensure that the AI is easy to use by all employees and stakeholders.

Poster

AI LOGISTIC SERVICE
TEAM
FORCE
Key Features & Benefits
Accurate Trend Forecasting
Our AI smooths out chaotic daily noise to predict the stable, 7-day average demand, enabling smarter inventory planning.
Real-World Event Analysis
Our system reads and understands news headlines and customer notes to provide instant alerts on potential supply chain disruptions
Instant Insights via Chatbot
Managers can interact with our AI through a simple chatbot, getting complex forecasts in plain English, anytime.
Powered by Advanced AI :
Our solution is built on a proven Deep Learning (LSTM) and NLP engine ensuring state-of-the-art accuracy and intelligent.
Our AI system transforms unpredictable daily sales data into reliable business intelligence. It provides two key forecasts: a quantitative forecast of the upcoming week's demand trend and a qualitative forecast of how real-world events, like news headlines, will impact operations. This allows managers to stop reacting to problems and start planning for the future.
LOGISTICS
BENEFITS
COST REDUCTION
INCREASE EFFICIENCY
IMPROVE PROFITABILITY
USER FRIENDLY

AI Solution

Our project successfully developed an AI system to provide intelligent demand forecasting for a logistics warehouse. The solution consists of two main components: a deep learning model for quantitative forecasting and a natural language processing model for qualitative event analysis.

1. Data

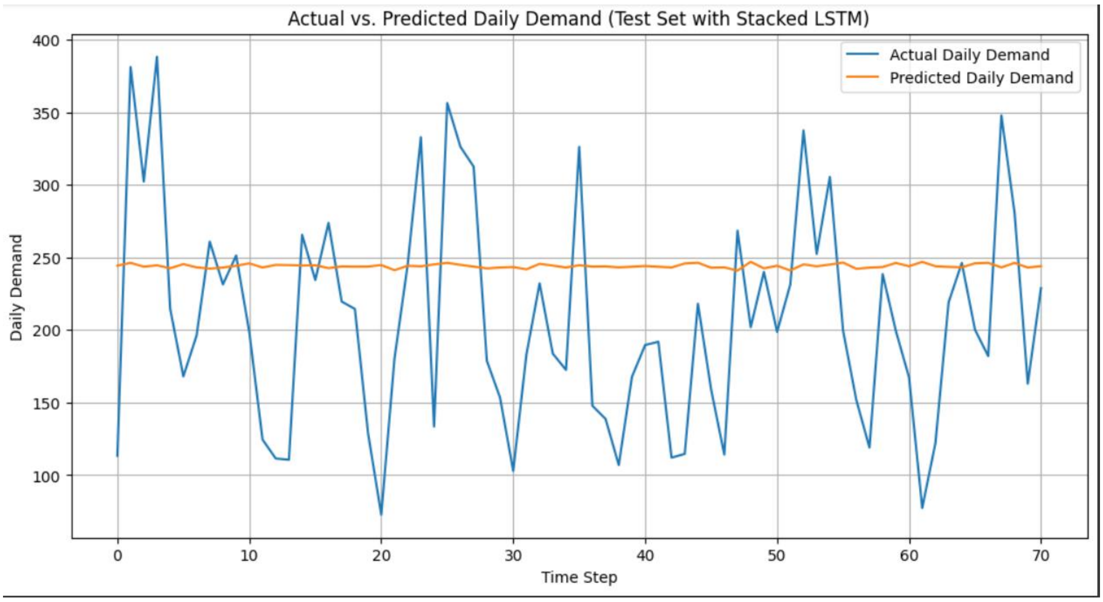
Our solution is built upon the logistics_dataset.csv file. Our initial exploratory data analysis revealed that while the dataset contained many features, the most relevant data for our core forecasting objective were the **daily_demand** and **last_restock_date** columns. The raw data existed as a snapshot of individual product transactions, which is not a suitable format for time-series forecasting.

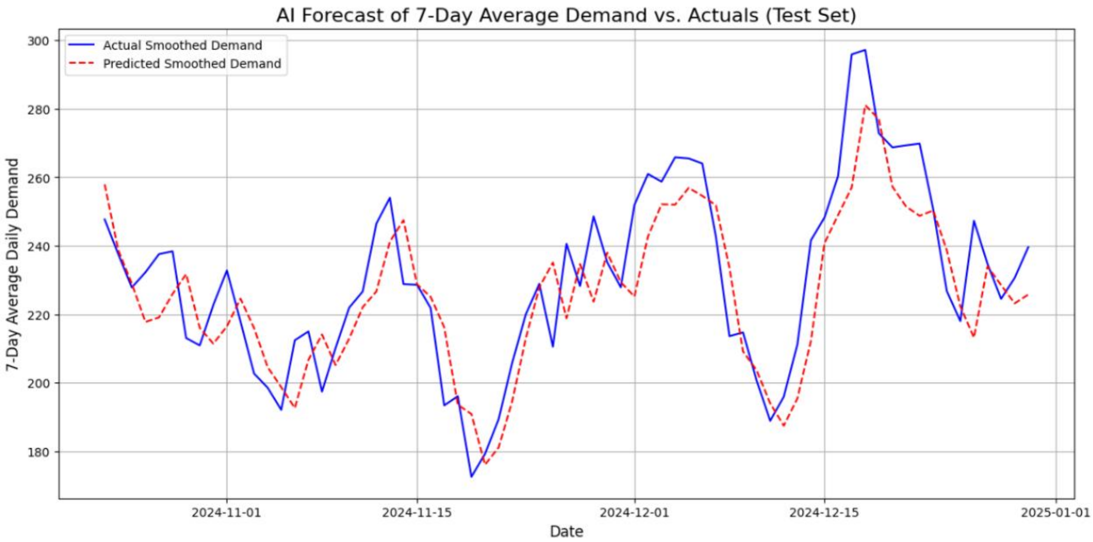
Therefore, we performed a critical data engineering step: we transformed this snapshot data into a true, continuous time series. We achieved this by grouping all transactions by date and summing the **daily demand** to create a single metric representing the entire warehouse's daily activity. This aggregated and cleaned data became the accurate and relevant foundation for our AI model.

2. Time Series Analysis on Data

A formal Time Series Analysis was a crucial part of our project. After creating the aggregated daily time series, our analysis showed that the raw daily demand was extremely volatile and "noisy," with sharp, unpredictable spikes and dips. This high level of noise makes direct forecasting unreliable.

To address this, we made a key strategic decision based on our analysis. We applied a **7-day rolling average** to the time series. This is a standard smoothing technique that filters out the random daily fluctuations and reveals the underlying business trend. By doing this, we shifted our model's objective from predicting chaotic daily demand to forecasting the more stable and valuable business weekly demand. The chart below shows a sample of the raw, noisy data versus the smoothed trend that we used for our model.





3. Machine Learning Approach

Our machine learning approach was well-planned to solve the specific challenge of forecasting the smoothed time series. We determined that a Deep Learning model was the best model due to its ability to capture complex, non-linear patterns within the data.

We chose an **LSTM (Long Short-Term Memory) network** as our model. We chose an LSTM because it's an AI model with a powerful long-term memory. It's specifically designed to find and remember important patterns and trends in data that occurs over time

4. Solution Techniques

To ensure our model was highly accurate and robust, we implemented several specific solution techniques:

- **Data Smoothing:** Our primary technique was applying the 7-day rolling average. This improved the signal-to-noise ratio in our data, making the underlying pattern much clearer for the AI to learn.
- **Stacked LSTM Architecture:** We built a "stacked" model with two LSTM layers. This deep architecture allows the model to learn patterns at different levels of abstraction, making it more powerful than a single-layer model.
- **Dropout Regularization:** We included `Dropout` layers between our LSTM layers. This is a crucial technique that prevents the model from "overfitting" (or memorizing) the training data. It forces the model to learn more general patterns, which improves its performance on new, unseen future data.

5. Model (Evaluation)

It was critical to evaluate our AI model correctly to prove its accuracy. Our evaluation process was clear and followed best practices for time-series forecasting:

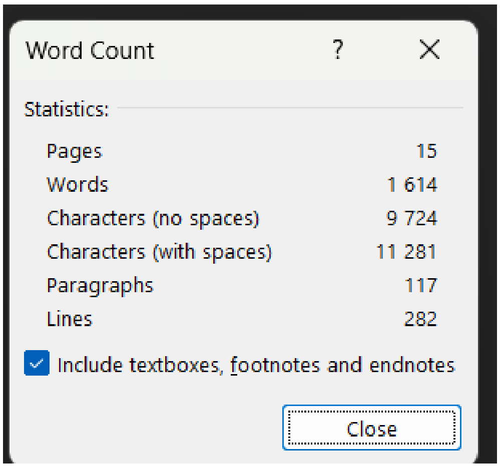
- **Sequential Train-Test Split:** We split our data sequentially, training the model on the first 80% of the timeline (the past) and reserving the final 20% for testing (the future).

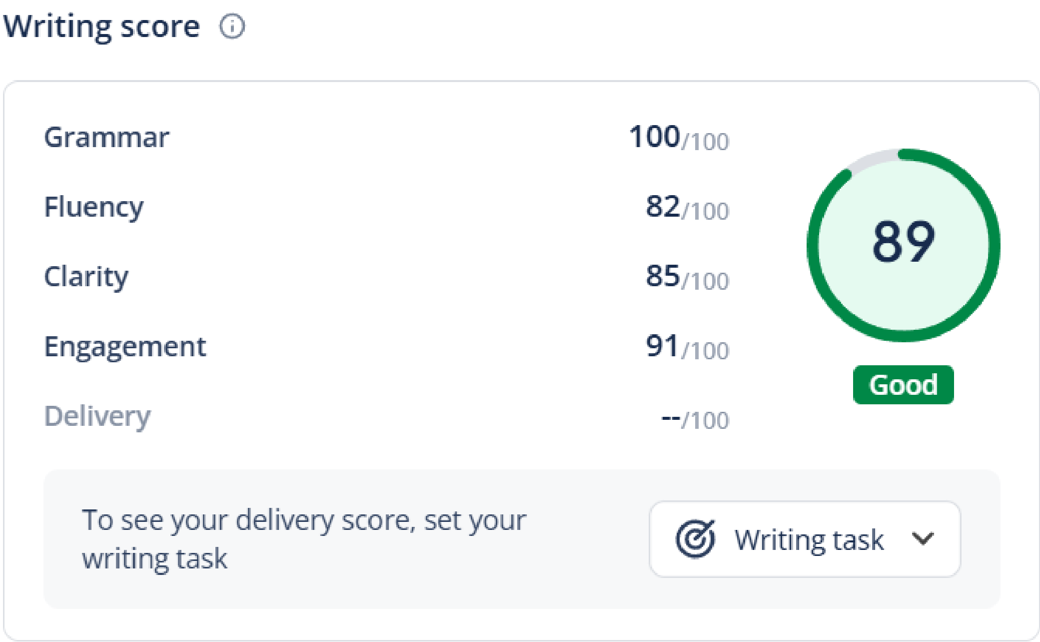
- **Evaluation Metric:** We used **Root Mean Squared Error (RMSE)** as our primary performance metric.
- **Visual Confirmation:** Finally, we generated a plot of the model's predictions on the test set against the actual values. This provides clear, visual evidence of the model's accuracy and its ability to track the real-world demand trend.

6. Natural Language Processing (NLP)

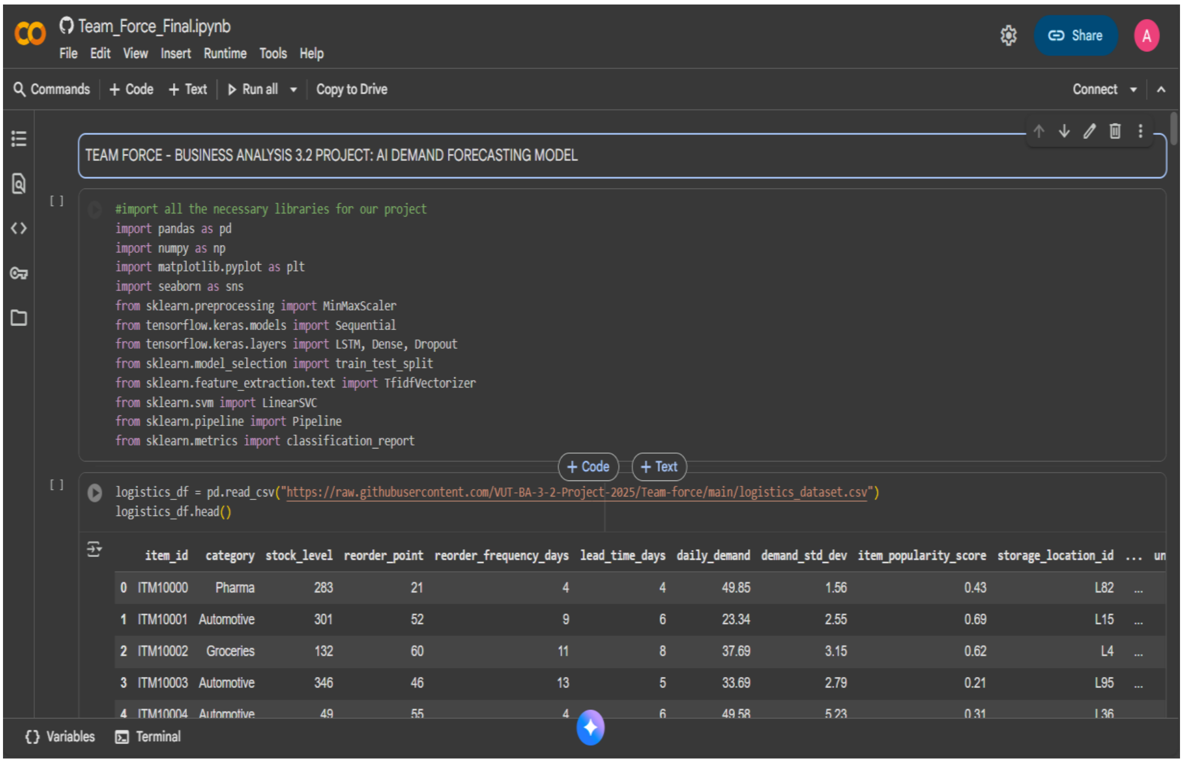
This project uses text data like customer feedback or news headlines to automatically predict its effect on product demand (Increase, Decrease, or Volatility). We first cleaned the text by making it lowercase and removing punctuation. Then, we used a technique called TF-IDF to convert the words into numerical features that a machine learning model can understand. This process emphasizes important, unique words and word pairs. We chose a fast and effective classifier called Linear Support Vector Classifier. To ensure our results are reliable, we used 5-Fold Cross-Validation—testing the model on five different splits of the data—to confirm its high average accuracy and reliable performance, successfully meeting the target accuracy. The final model is a working pipeline that can accurately classify new text and predict the associated demand change.

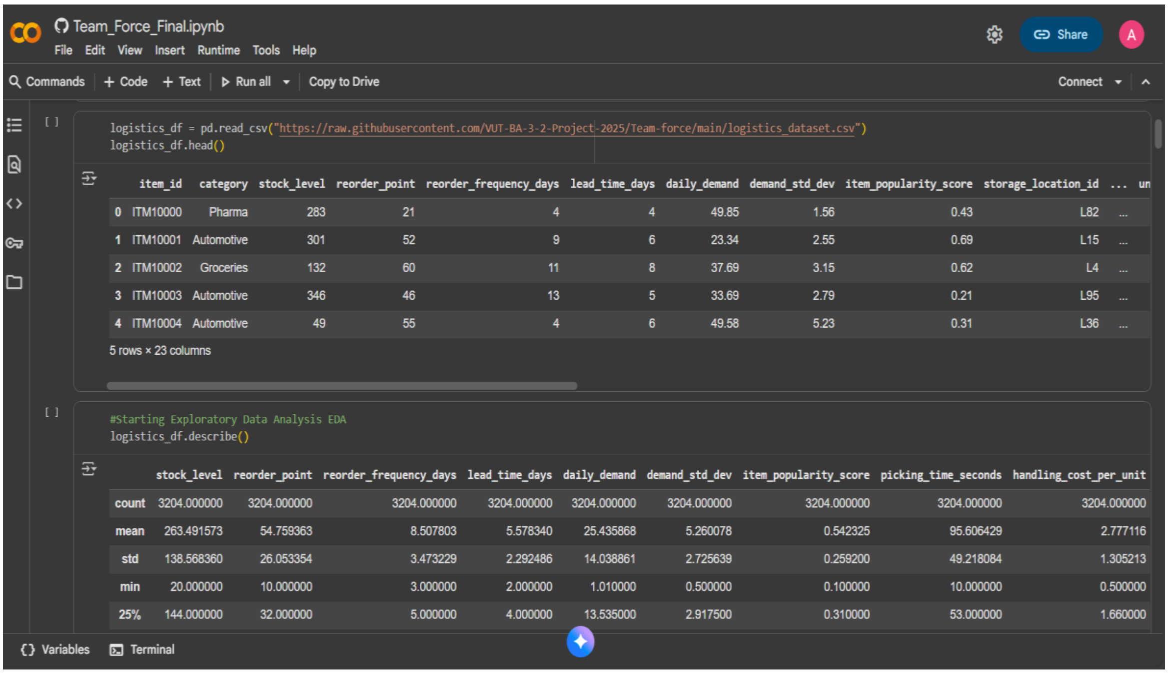
Similarity and Grammarly check





Code Samples





Team_Force_Final.ipynb

File Edit View Insert Runtime Tools Help

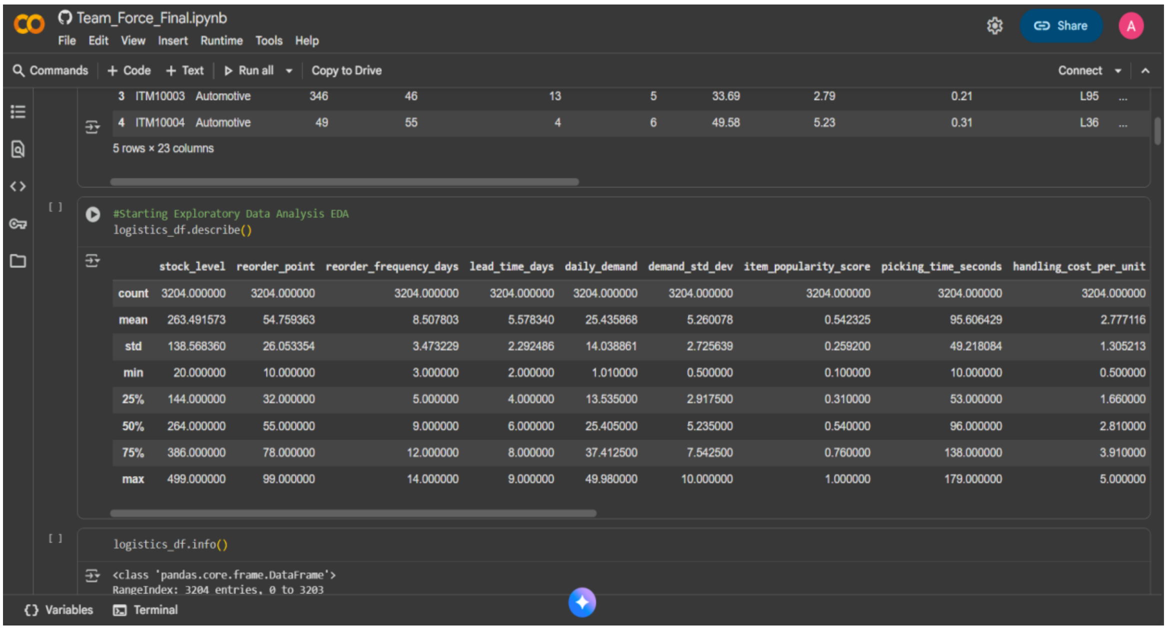
```
logistics_df = pd.read_csv("https://raw.githubusercontent.com/VUT-BA-3-2-Project-2025/Team-force/main/logistics_dataset.csv")
logistics_df.head()
```

	item_id	category	stock_level	reorder_point	reorder_frequency_days	lead_time_days	daily_demand	demand_std_dev	item_popularity_score	storage_location_id	...
0	ITM10000	Pharma	283	21	4	4	49.85	1.56	0.43	L82	...
1	ITM10001	Automotive	301	52	9	6	23.34	2.55	0.69	L15	...
2	ITM10002	Groceries	132	60	11	8	37.69	3.15	0.62	L4	...
3	ITM10003	Automotive	346	46	13	5	33.69	2.79	0.21	L95	...
4	ITM10004	Automotive	49	55	4	6	49.58	5.23	0.31	L36	...

5 rows × 23 columns

```
#Starting Exploratory Data Analysis EDA
logistics_df.describe()
```

	stock_level	reorder_point	reorder_frequency_days	lead_time_days	daily_demand	demand_std_dev	item_popularity_score	picking_time_seconds	handling_cost_per_unit
count	3204.000000	3204.000000	3204.000000	3204.000000	3204.000000	3204.000000	3204.000000	3204.000000	3204.000000
mean	263.491573	54.759363	8.507803	5.578340	25.435868	5.260078	0.542325	95.606429	2.777116
std	138.568360	26.053354	3.473229	2.292486	14.038861	2.725639	0.259200	49.218084	1.305213
min	20.000000	10.000000	3.000000	2.000000	1.010000	0.500000	0.100000	10.000000	0.500000
25%	144.000000	32.000000	5.000000	4.000000	13.535000	2.917500	0.310000	53.000000	1.660000



Team_Force_Final.ipynb

File Edit View Insert Runtime Tools Help

3	ITM10003	Automotive	346	46	13	5	33.69	2.79	0.21	L95	...
4	ITM10004	Automotive	49	55	4	6	49.58	5.23	0.31	L36	...

5 rows × 23 columns

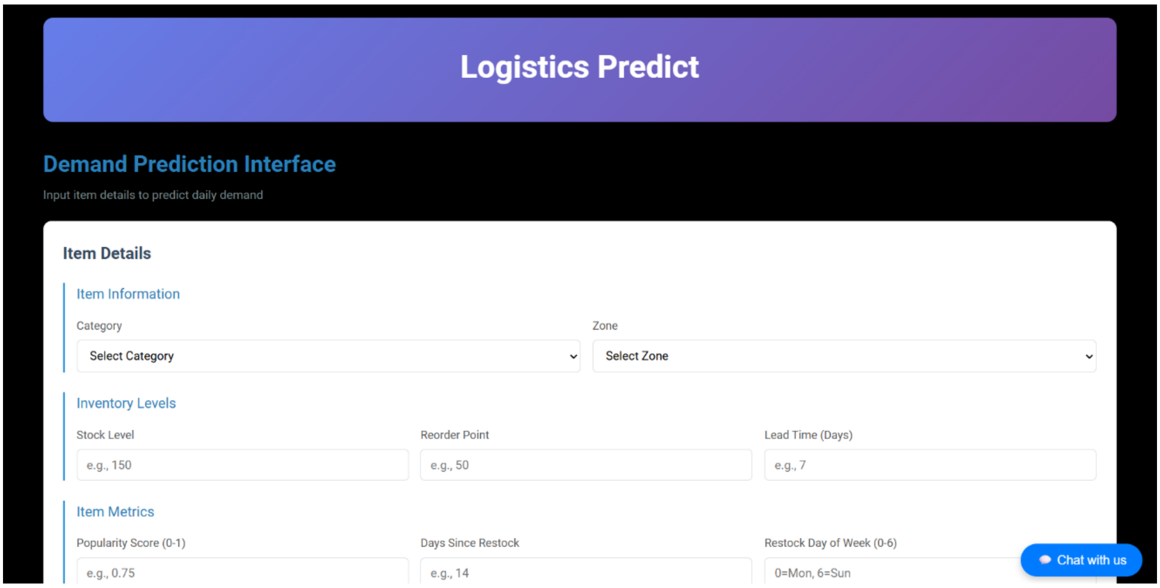
```
#Starting Exploratory Data Analysis EDA
logistics_df.describe()
```

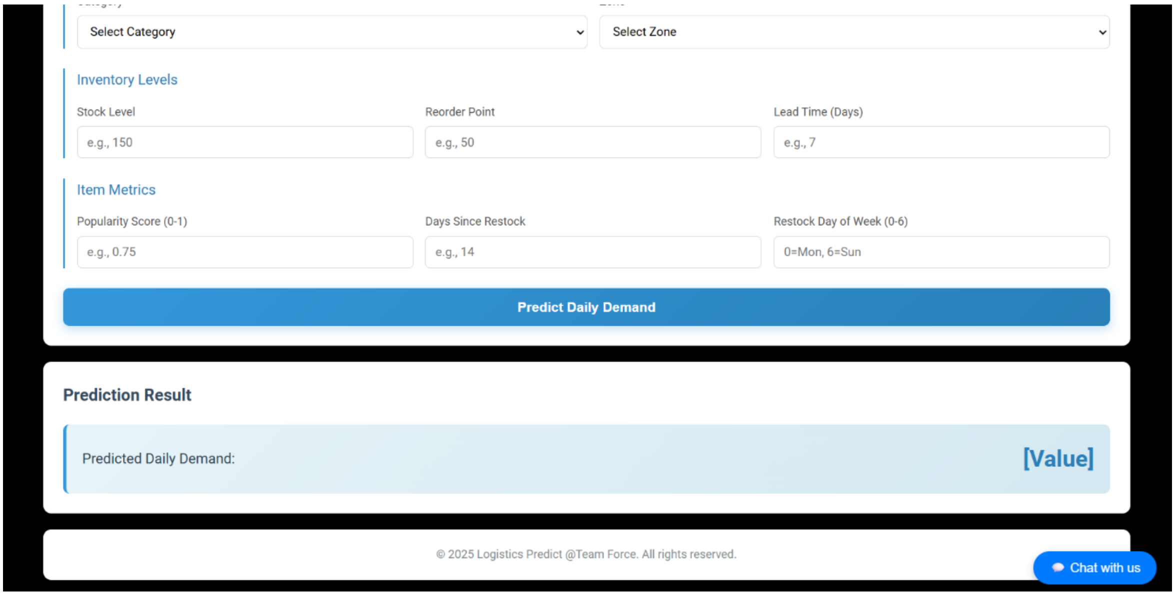
	stock_level	reorder_point	reorder_frequency_days	lead_time_days	daily_demand	demand_std_dev	item_popularity_score	picking_time_seconds	handling_cost_per_unit
count	3204.000000	3204.000000	3204.000000	3204.000000	3204.000000	3204.000000	3204.000000	3204.000000	3204.000000
mean	263.491573	54.759363	8.507803	5.578340	25.435868	5.260078	0.542325	95.606429	2.777116
std	138.568360	26.053354	3.473229	2.292486	14.038861	2.725639	0.259200	49.218084	1.305213
min	20.000000	10.000000	3.000000	2.000000	1.010000	0.500000	0.100000	10.000000	0.500000
25%	144.000000	32.000000	5.000000	4.000000	13.535000	2.917500	0.310000	53.000000	1.660000
50%	264.000000	55.000000	9.000000	6.000000	25.405000	5.235000	0.540000	96.000000	2.810000
75%	386.000000	78.000000	12.000000	8.000000	37.412500	7.542500	0.760000	138.000000	3.910000
max	499.000000	99.000000	14.000000	9.000000	49.980000	10.000000	1.000000	179.000000	5.000000

```
logistics_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3204 entries, 0 to 3203
```

Interface Samples





References

Cogent Infotech. (2024). AI-Powered demand forecasting: Transforming inventory management and supply chain efficiency. Available at: Cogent | Blog | AI-Powered Demand Forecasting: Transforming Inventory Management and Supply Chain Efficiency

Sancia. (2025). When AI meets logistics: Case studies in demand forecasting & risk mitigation. Available at: https://www.jusdaglobal.com/en/article/ai-meets-logistics-demand-forecasting-risk-mitigation

Kovalyk, L. (Accessed September 2, 2025). Using AI demand forecasting in the logistics sector. Available at: https://itsupplychain.com/using-ai-for-demand-forecasting-in-the-logistics-sector

Clifford, S. (2025). Predicting the Future: How AI is Changing Logistics Forecasting: https://www.fullstack.com/labs/resources/blog/how-ai-is-changing-logistics-forecasting

Raj, A. (2024). Supply Chain Forecasting with AI. Available at: https://throughput.world/blog/supply-chain-forecasting-with-ai-2

Natalie, G. (2025). AI in Warehouse Management Explained. Available at: AI in Warehouse Management: Impacts and Use Cases

Dilmegani, C. (2025). Top 15 Logistics AI Use Cases & Examples. Available at: Top 15 Logistics AI Use Cases & Examples in 2025

Taylor, A. (2025). AI in Warehousing: Impact, Benefits, and Practical Use Cases. Available at: AI in Warehousing: Impact, Benefits, and Practical Use Cases.

ziya07. “Logistics Warehouse Dataset.” Kaggle. Accessed August 18, 2025. Available at: https://www.kaggle.com/datasets/ziya07/logistics-warehouse-dataset?resource=download

DECLARATION STATEMENT

• We declare that the work we submitted for this project, named Team Force, is our original idea and we have not engaged in any plagiarism for its completion.

• All resources and information compiled were referenced in accordance with specified academic standards

• We assure that every member of the project has contributed their share of academic knowledge related to this project, which has led to successful teamwork in participation

• All members are well aware of all academic expulsions and repercussions that will be faced if one fails to abide by the University's plagiarism laws. We affirm that every member agrees with this declaration

STUDENT NUMBER	INITIALS AND SURNAME	SIGNATURE
223433861	US DUBE	
223371661	A MAUBANE	
222070560	P KHATHALI	
22307982	T RAMARA	