Demand Prediction Project – Summary Report

1. Project Overview

This project implements a machine learning demand forecasting system using multiple modeling techniques to predict future sales for various grocery items (Apple, Banana, Tomato, Potato, Spinach).
The solution includes:

A Streamlit dashboard for interactive visualization and prediction.

Four different forecasting models:

- Moving Average
- Linear Regression
- Time Series (Holt-Winters)
- XGBoost (primary model)

A comparison framework to evaluate model performance across items.
Github: https://github.com/Calxy72/demand-prediction-xgboost
Presentation video: gromally assignment.mov

2. Technical Implementation

2.1 Data Pipeline

Historical sales data loaded from enhanced_sales_data.csv

Feature engineering:

- Lag features (sales from previous days)
- Rolling statistics (mean, std)
- Time features (day of week, weekend flag)

2.2 Model Training

- XGBoost: Trained with engineered features, optimized for time-series prediction.

- Baseline models: Implemented as reference points (Moving Average, Linear Regression, Holt-Winters).

2.3 Evaluation

- Metrics: MAE, RMSE, R^2
- Test set: Last 7 days of available data
- Visualization: Plotly charts for interactive exploration

3 Predictions Insight

- XGBoost: Stable, realistic predictions with minimal fluctuation
- Linear Regression: Flat predictions (fails to capture patterns)
- Time Series: Captures seasonality but high variance
- Moving Average: Simple but volatile predictions

4. Conclusion

- The project delivers a complete, production-ready demand forecasting system:
- Robust Backend: FastAPI serving multiple ML models
- User-Friendly Interface: Interactive Streamlit dashboard
- Integration Ready: REST API for system-to-system communication
- Proven Accuracy: XGBoout outperforms traditional methods
- Scalable Architecture: Ready for cloud deployment

Assistance used:
ChatGPT, Deepseek.