Final Project Documentation

# Project Title: Smart Perishable Demand Forecasting in Grocery Retail

## 1. Executive Summary

This project focused on solving a common challenge in grocery retail—forecasting demand for perishable items like fruits, vegetables, milk, and meat. By integrating weather data, calendar effects, and product/store-level details, a LightGBM-based model was built to predict daily sales. A user input script and a Tableau dashboard were created to make the model actionable and visually insightful.

## 2. Problem Statement

Grocery retailers often struggle with unpredictable demand for perishables. Overstocking leads to spoilage and financial loss, while understocking causes missed sales and unhappy customers. External variables such as weather, day-of-week effects, and store dynamics compound the challenge.

## 3. Objectives

* Build a daily demand forecasting model at the SKU-store level.
* Include calendar, weather, and store-level features to enhance accuracy.
* Support smarter inventory planning decisions in grocery retail.
* Provide actionable insights through a Tableau dashboard and prediction tool.

## 4. Data Sources

The dataset (Data\_for\_2024.csv) includes historical sales data, temperature, rainfall, product type, store ID, and date features. The data was cleaned, formatted, and prepared for modeling using pandas in Python.

## 5. Methodology

* Loaded and explored the dataset using pandas (`head`, `info`, `describe`, `isnull`).
* Converted the date column to datetime format.
* Checked for outliers in temperature and rainfall using boxplots.
* Filtered extreme values and handled missing data.
* Trained a LightGBM regression model to predict units sold.
* Evaluated model performance using MAE and RMSE.
* Built a user input script to simulate store-level predictions.
* Created a Tableau dashboard to visualize demand trends and insights.

## 6. Key Findings and Insights

The model achieved a low MAE and RMSE, indicating strong predictive performance. The insights showed clear seasonality, weather-based variation, and product-type-driven trends in demand. Store-level behavior varied significantly, reinforcing the need for localized forecasts.

## 7. Business Impact

* Reduce spoilage costs by avoiding overstocking.
* Improve product availability through more accurate forecasting.
* Optimize inventory holding costs and shelf space allocation.
* Provide daily actionable forecasts for retail managers.

## 8. Limitations

The model currently does not account for promotional campaigns, pricing effects, or supplier constraints. It is based on historical weather and sales, and may not perform well under extreme/unusual conditions.

## 9. Future Scope

* Incorporate promotional and pricing data into the model.
* Deploy model as a web app or integrate with store ERP systems.
* Use real-time weather APIs for live forecasting.
* Expand to multi-store, multi-region scaling.

## 10. Appendices

* Python code (`code.ipynb`) including model training and prediction script.
* Original dataset (`Data\_for\_2024.csv`) used for development.
* Project SOW document for formal scoping and planning.
* Tableau workbook (`Tableau.twb`) for visual analytics.