

Sales-Inventory-Shrinkage-Analysis-and-Prediction

By

Venkata Sai Dhanush MIRIYALA

Inventory Shrinkage Analysis & Prediction

This project analyzes retail inventory shrinkage by combining datasets on stock, sales, and returns to identify product loss patterns. Using data cleaning, aggregation, and exploratory analysis, it builds predictive models (Linear Regression, Random Forest, Gradient Boosting) to forecast shrinkage. The project includes interactive visualizations and a dashboard to help businesses monitor inventory loss, improve operational efficiency, and reduce financial impact.

Dataset Description

Three main datasets are used in this project:

1. ****Inventory Dataset****

Contains daily records of stock movements for products.

- `date`: Date of record
- `product_id`: Unique product identifier
- `stock_in`: Number of units added to inventory
- `stock_out`: Number of units removed or sold from inventory
- `stock_remaining`: Units remaining at day end

2. ****Sales Dataset****

Tracks daily sales for each product.

- `date`: Date of sale
- `product_id`: Unique product identifier
- `quantity_sold`: Units sold
- `unit_price`: Price per unit

3. ****Returns Dataset****

Records returned products and reasons.

- `date`: Date of return
- `product_id`: Unique product identifier
- `quantity_returned`: Units returned
- `return_reason`: Reason for return (e.g., damaged, wrong item)

Methodology & Pipeline

1. **Data Cleaning & Preparation**

- Loaded datasets, ensured consistent date formats.
- Aggregated total stock added (`stock_in`), sales (`quantity_sold`), and returns (`quantity_returned`) by `product_id`.
- Merged datasets to calculate **shrinkage** as:
$$\text{shrinkage} = \text{stock_in} - (\text{quantity_sold} - \text{quantity_returned})$$

2. **Exploratory Data Analysis (EDA)**

- Visualized top products by shrinkage using bar charts.
- Calculated KPIs: total sales revenue, total returns, shrinkage per product.

3. **Model Building**

- Used Linear Regression to predict shrinkage based on:
 - Stock added
 - Quantity sold
 - Quantity returned
- Split data into training and testing sets.
- Evaluated model performance using R^2 score (~ 0.99).

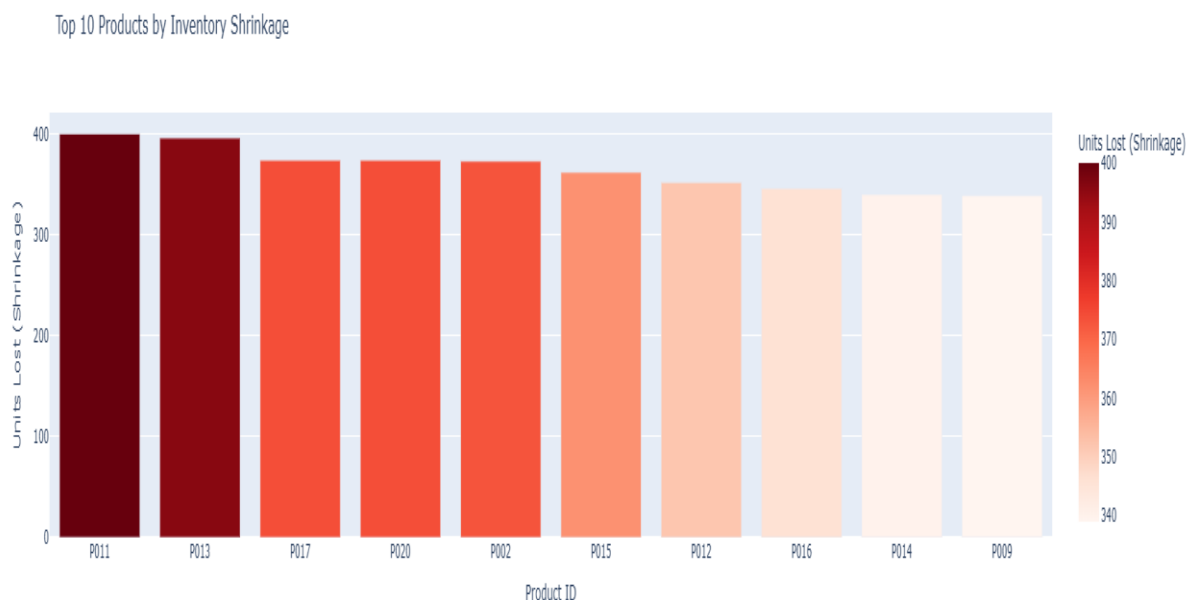
4. **Prediction on New Data**

- Loaded new monthly inventory and sales data.
- Predicted shrinkage values using the trained model.
- Saved predictions for business use.
- Jupyter Notebook for interactive development and reporting

Technologies Used

- Python 3
- Pandas for data manipulation
- Matplotlib & Plotly for visualization
- Scikit-learn for regression modeling
- Jupyter Notebook for interactive development and repor

TOP 10 Products by Inventory shrinkage



Model

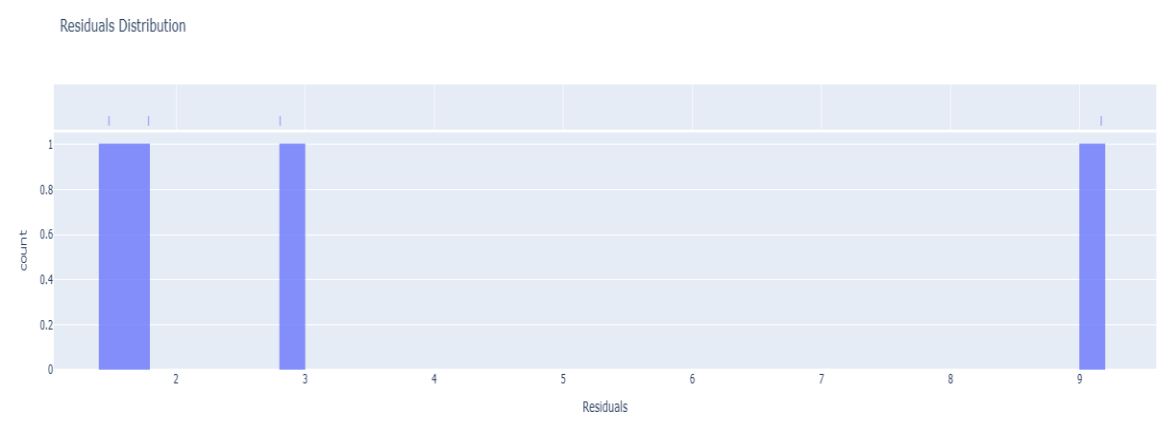
* Model R^2 score: 0.9895746689563556

* Random Forest R^2 : 0.7687706838820284

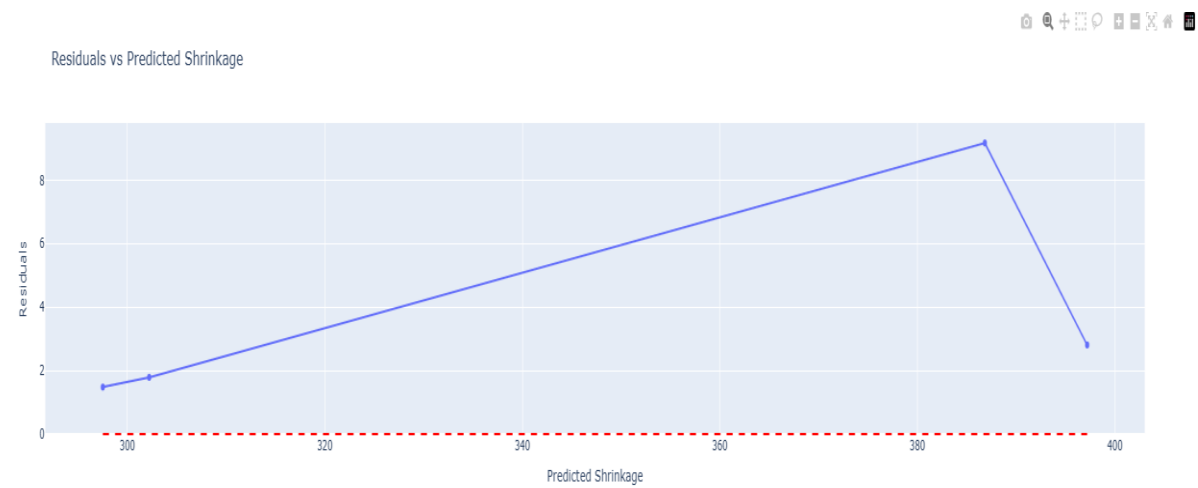
* Gradient Boosting R^2 : 0.8382298875721109

Linear regression is better in compare dto both Random Forest and Gradient Boosting

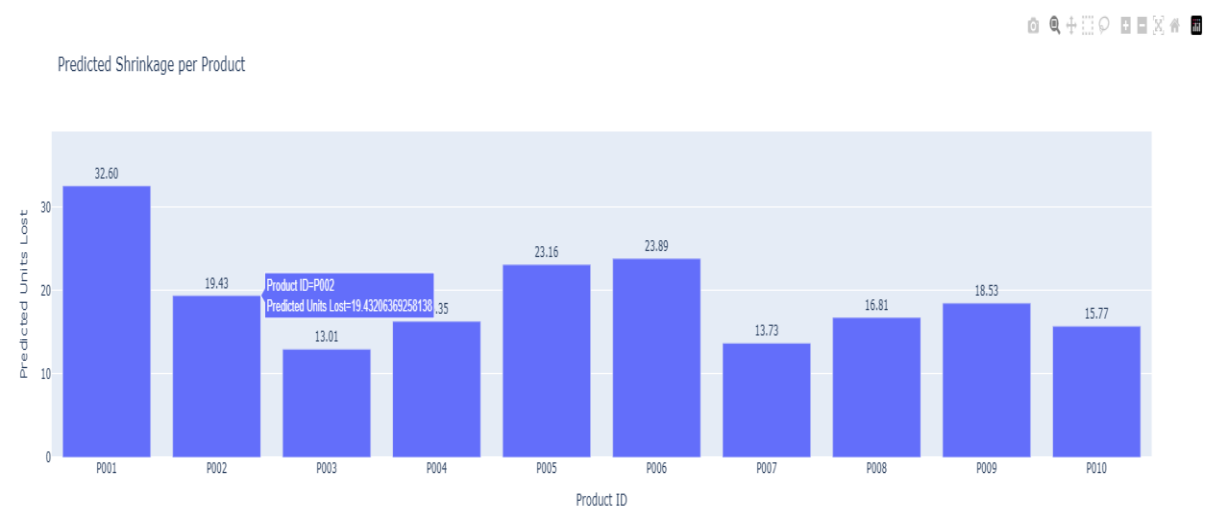
Residual Distribution



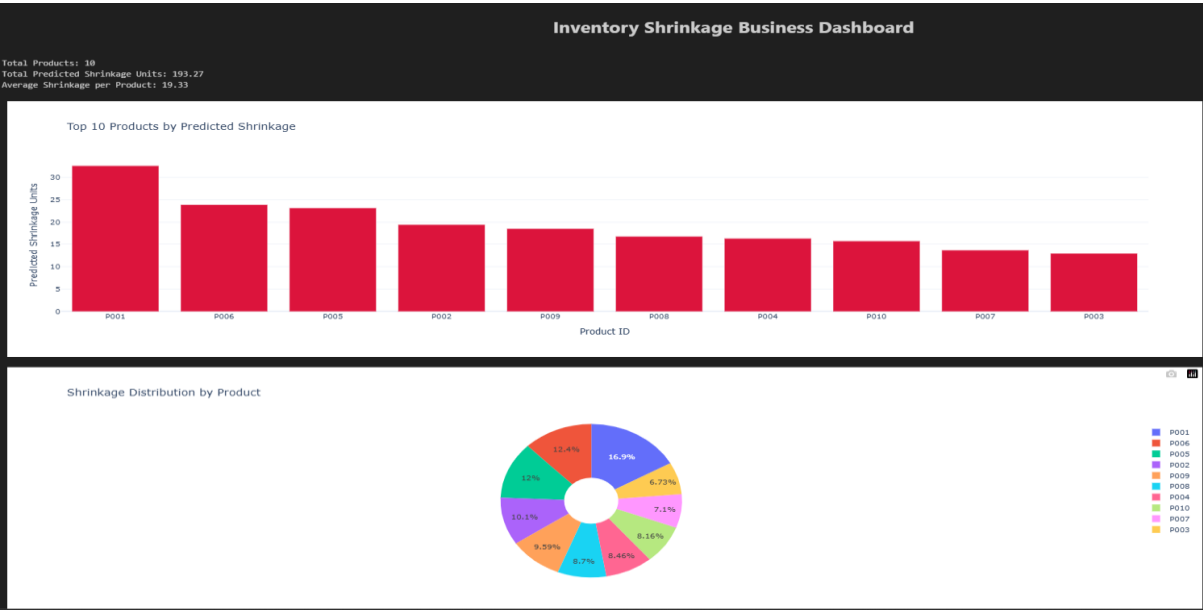
Residual vs Shrinkage



Predicted shrinkage per Product



Dashboard

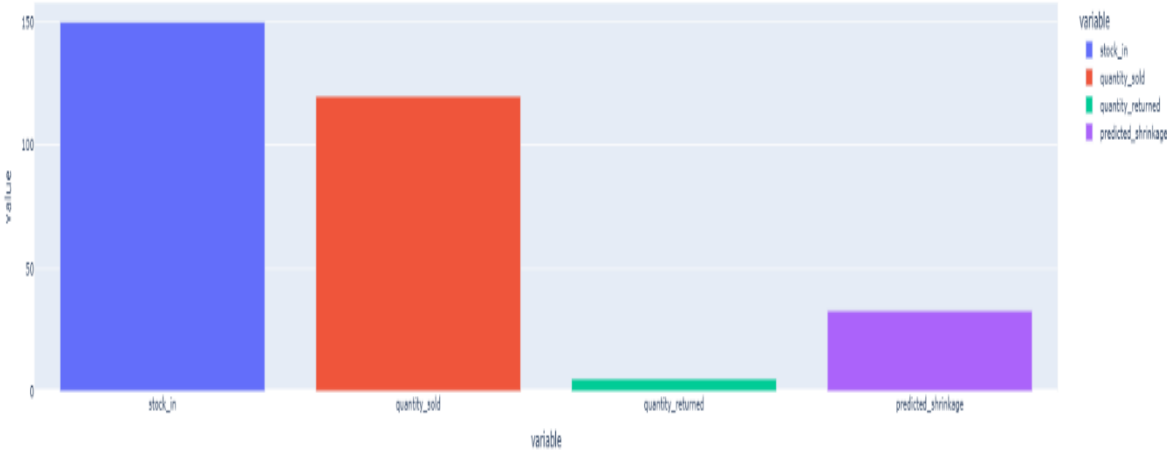


Inventory Overview

Product:



Inventory Overview for P001



Insights & Business Impact

- Top shrinkage products identified:** Helps prioritize inventory audits.
- Accurate shrinkage prediction:** Enables proactive stock management and loss reduction strategies.
- Returns analysis:** Provides insight into product quality or customer satisfaction issues.
- Interactive dashboards:** Facilitate easy monitoring of KPIs by business teams.

Conclusion

This project demonstrates how combining sales, inventory, and return data can reveal hidden losses through shrinkage analysis. Predictive modeling empowers retailers to anticipate losses and optimize stock management, ultimately improving profitability.