

Dillard's Predictive Analysis for Inventory Cost Improvement

Presented by: Ayush Agarwal, Kevin Li, Seth Kazarian, Yilin Shan
For: MLDS 400 Class Project, G11

Agenda

:

- Objective and Executive Summary
- Exploratory Data Analysis (EDA)
- Data Visualizations
- Feature Engineering
- Model Approach and Evaluation
- Business Impact
- Strategic Value
- Application and Challenges
- Conclusion and Future Directions

Objective and Executive Summary

- Develop a predictive model for forecasting inventory selling prices.
- Aim to align inventory levels with projected sales to reduce costs.
- Achieved a total reduction in inventory costs by \$117.2 million.
- Improvement of 21.6% over the baseline model across three quarters.

21.6%

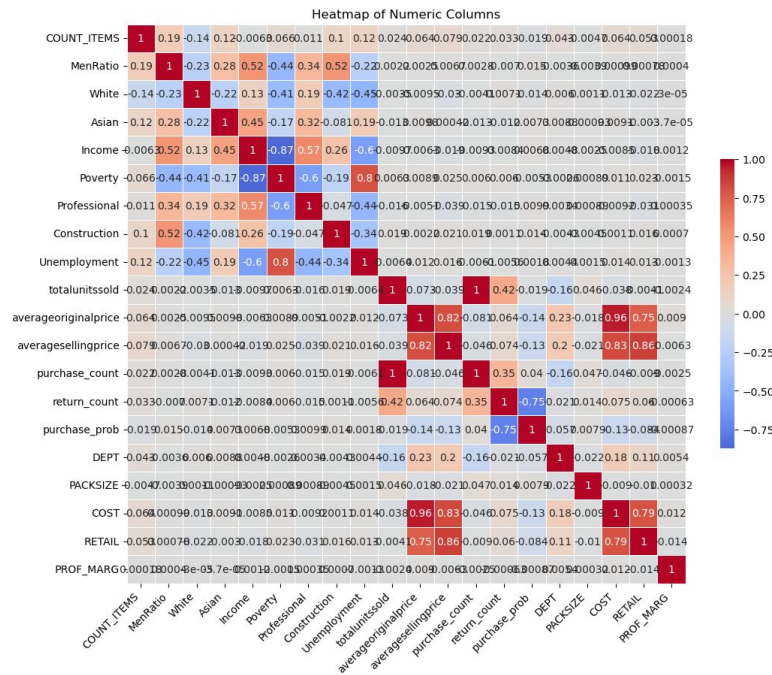
ROI Improvement Over Baseline

\$ 117 million

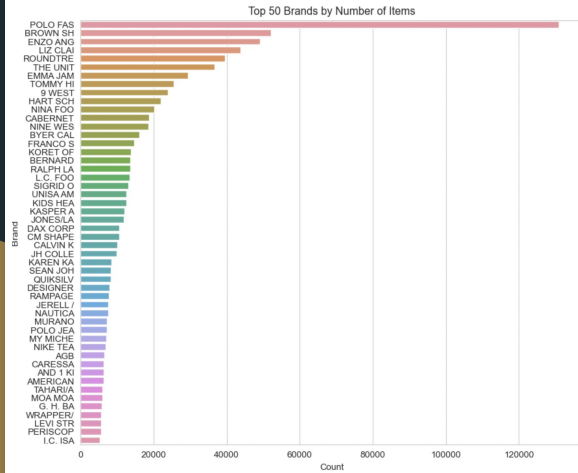
Dollar Saved in Inventory after
Labor Cost

Exploratory Data Analysis (EDA)

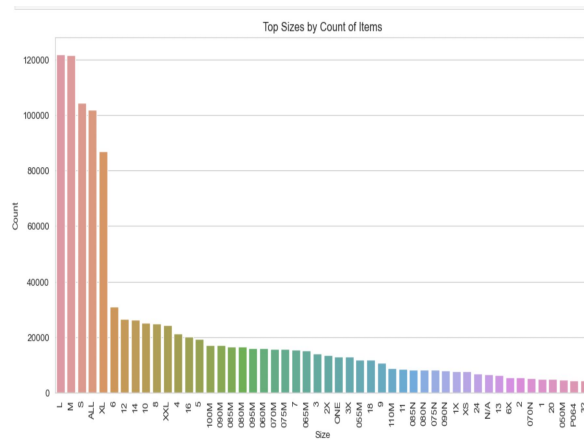
- Checked for types, duplicates and abnormal patterns and performed data cleansing by methods like imputations on missing values and removing extra spaces.
- Plotted histograms, bar plots, frequency tables, box plots to visualize and examined the distribution.
- Used z-scores and interquartile ranges for outlier removal to ensure data integrity.
- Analyzed SKUinfo table for correlation between attributes like brand and color with sales trends and inventory turnover.
- In the STRInfo table, included geographical spread, market penetration, and sales performance per region.
- Generated heatmap and observed insights:
 - a. High correlation among poverty, income and unemployment.
 - b. High correlation among cost, average selling price, average original price, and retail.



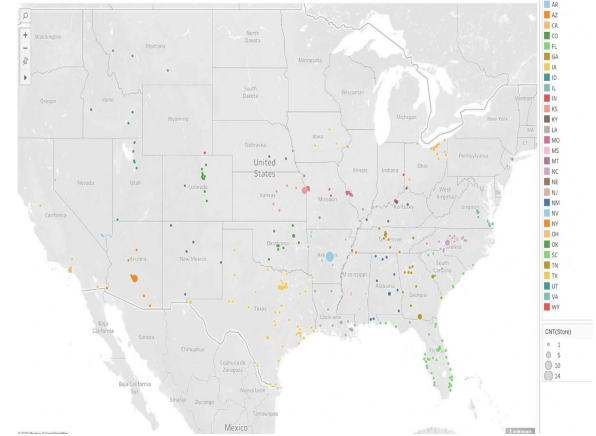
Data Visualizations



Top 50 Brands by Number of Items



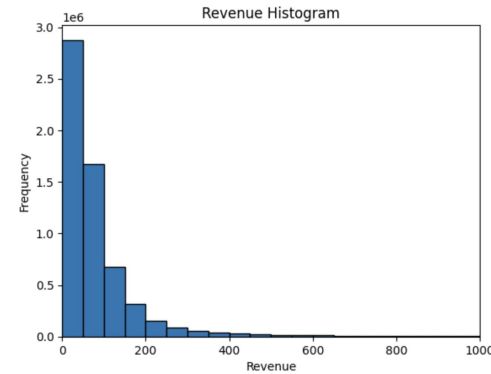
Top Sizes by Count of Items



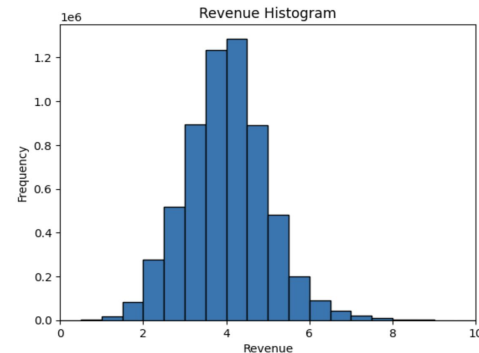
Geographical Store Distribution

Feature Engineering

- Grouped transaction table by SKU, STORE and Quarter to create a comprehensive dataframe which contain store level and sku level information at quarterly level.
- Engineered store-level demographic features like StockCount, Men Ratio, Poverty etc to predict store-specific revenue and inventory costs.
- Created dummy variables for product-level features such as brands, class id, states etc for inventory and promotions recommendations.
- Apply logarithmic transformation to cost related features like COST, Retail, average cost price and revenue.



Revenue Distribution before logarithmic transformation



Revenue Distribution after logarithmic transformation

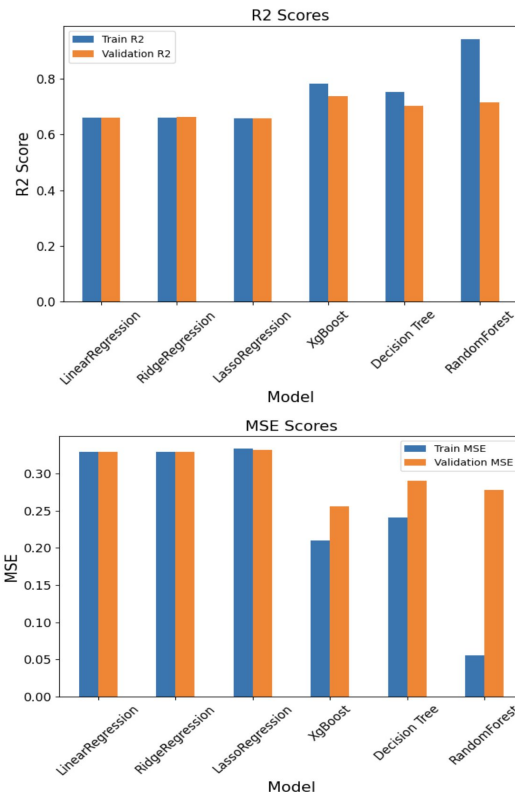
Modeling Approach and Evaluation

Experimentation

- Utilized Linear Regression, Ridge Regression, Lasso Regression, Decision Trees, Random Forest and XGBoost.
- K-Fold Cross Validation used to find best parameters for the algorithms.
- XGBoost Performs the best among all the models.

Evaluation of XGBoost

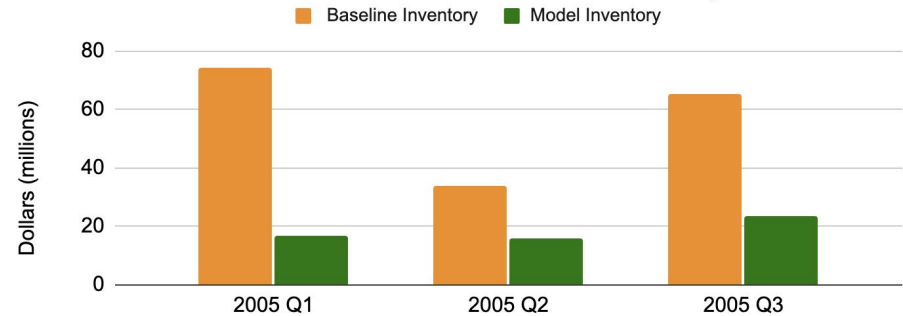
- Trained on a dataset of 320,000 rows and 1,205 variables.
- Validated on a subset of 80,000 rows.
- Test Scores
 - R2 Score = 73.72%
 - MSE Score = 0.2556



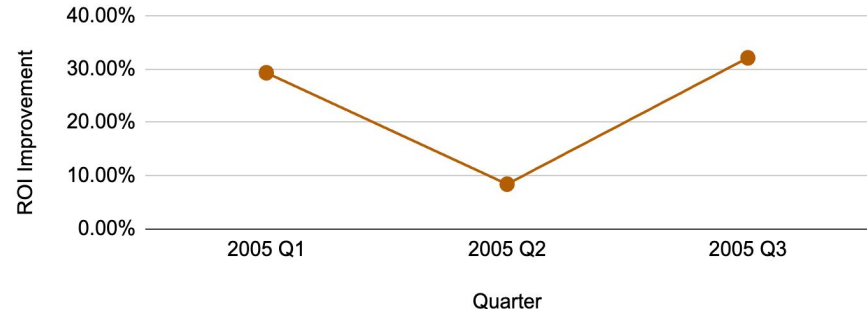
Business Impact

- The model resulted in a total reduction in inventory costs for overstocked items of \$117 million for Quarter 1 to Quarter 3 of 2005. We establish our baseline as sales from the previous quarter, and measure cost as
- Enhanced ROI significantly each quarter compared to the baseline.

Cost Incurred from Overstocked Inventory



ROI Percentage Improvement by Quarter



Strategic Value

- A predictive model sharpens decision-making, optimizes inventory, and improves customers engagement.
- Improving the model's ability to adapt to market trends, consumer behavior can provide deeper insight in revenue generation and waste reduction.
- Efficient inventory management reduces carbon footprint throughout the supply chain, reducing cost for Dillards and suppliers. .



[Reference](#)

Applications and Challenges

Applications

- Dillard can use our extensive EDA, feature engineering, and modelling techniques to train new model on data from recent years (2005 - 2023). This allows them to discover new insights on product, store, and transaction trends and streamline the model building process.
- By using the model to predict 2024 Q1 sales and onward, Dillard can monitor inventory levels accurately, reducing probability of severely overstocked items.

Challenges

- Due to external market variables and consumer trends, the model's revenue prediction model still has inherent errors that need enhancement.

Conclusion and Future Directions

- **Successful Implementation of Predictive Analytics:** The project effectively utilized exploratory data analysis feature engineering to build a predictive model that accurately forecasts of inventory selling prices. This process included analyzing various data points and extracting key insights, which contributed significantly to aligning inventory levels with projected sales, and ultimately reducing inventory costs by \$117.2 million, a 21.6% improvement over the baseline model across three quarters.
- **Future Directions for Enhanced Accuracy and Efficiency:** The project underscores the potential for integrating more advanced techniques like time series forecasting, deep learning, and transformer-based models in the future. These enhancements aim to further refine the accuracy of predictions and adapt to evolving market trends and consumer behaviors, thereby continuing to improve financial efficiency and inventory management at Dillard's.



Thank You