**BigMart Sales Prediction Analysis**

**Objective**

Develop a predictive model for BigMart outlet sales using historical transaction data and product attributes.

**Data Preprocessing**

Combined training and test datasets for consistent preprocessing. Key steps included:

* Standardized inconsistent fat content labels
* Imputed missing item weights using identifier-based grouping
* Filled outlet size gaps using outlet type distributions
* Corrected zero visibility values with median substitution

**Feature Development**

Created additional features:

* Outlet age from establishment year
* Item categorization from identifier patterns
* Price-to-weight ratios
* Relative visibility metrics
* Item frequency indicators
* Price tier binning
* Cross-feature interactions

**Modeling:**

Implemented three algorithms with performance comparison:

* Random Forest
* Gradient Boosting
* Weighted ensemble combining both models

**Target Encoding Implementation**

Applied outlet-item performance encoding using training data exclusively to prevent information leakage.

**Model Selection**

Automated selection based on validation RMSE performance across all three approaches.

**Results:**

Achieved a best score of 1152.1533214187 on the Submission portal.