'E-Commerce Market Analysis and Insights

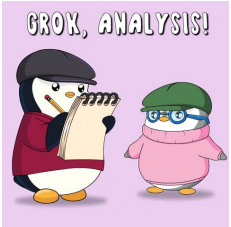
Team Members:

Shreeram Venkatesh - A20519833

Santhosh Mani - A20518627

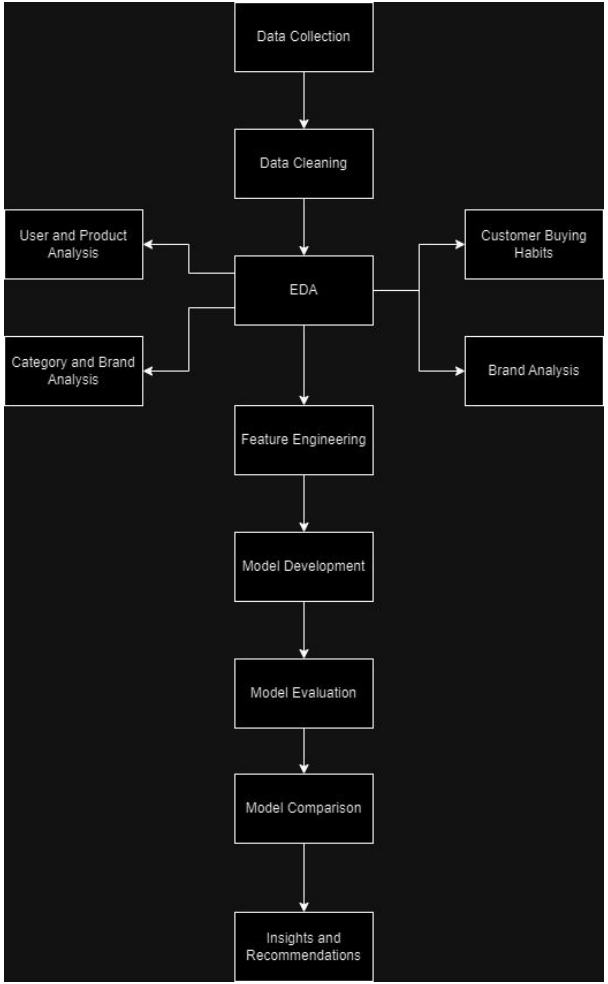
Gautham Kumar Ramappa Shanmugam - A20512380

Executive Summary of E-commerce Analysis



- Comprehensive analysis of e-commerce customer behaviors and market trends.
- Identification of key consumer behavior trends.
- Insights on product category and brand popularity.
- Recommendations for enhancing user engagement and conversion rates.
  -

Methodology Overview



Data Collection and Cleaning

1. Data Collection

   Sourced from Kaggle: E-Commerce Sales Dataset and Unlock Profits with E-Commerce Sales Data.

   Included user interactions, product details, and transaction data.

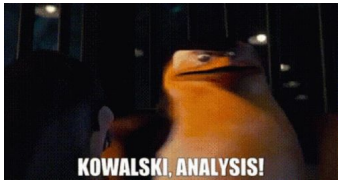
2. Data Cleaning

   Standardization of formats.

   Addressing missing values.

   Outlier detection and treatment.

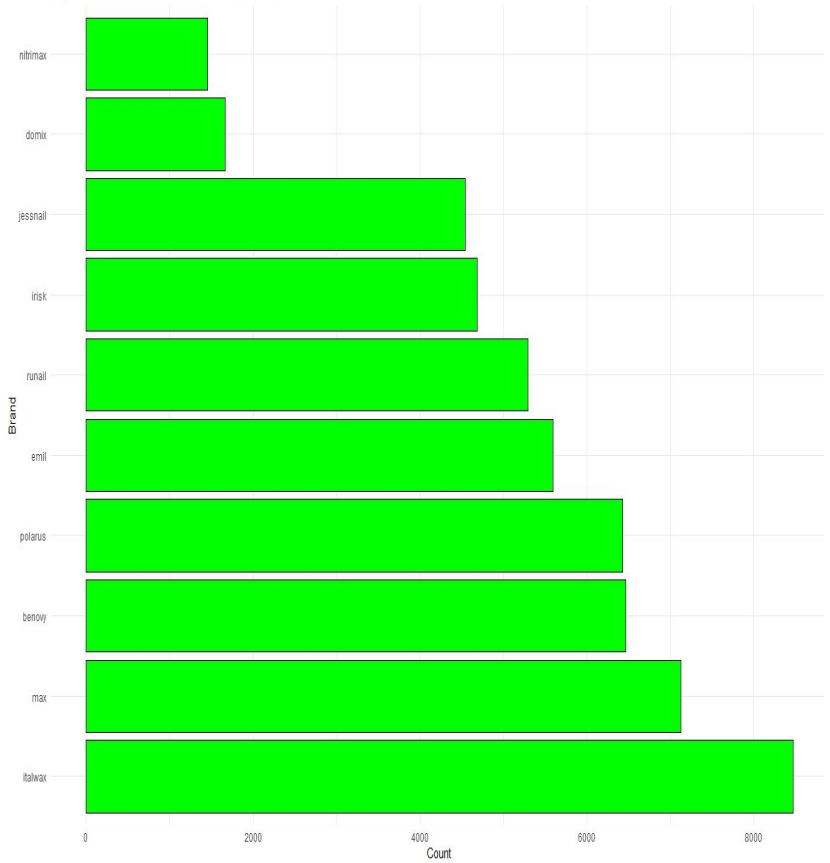
Exploratory Data Analysis - User and Product Analysis



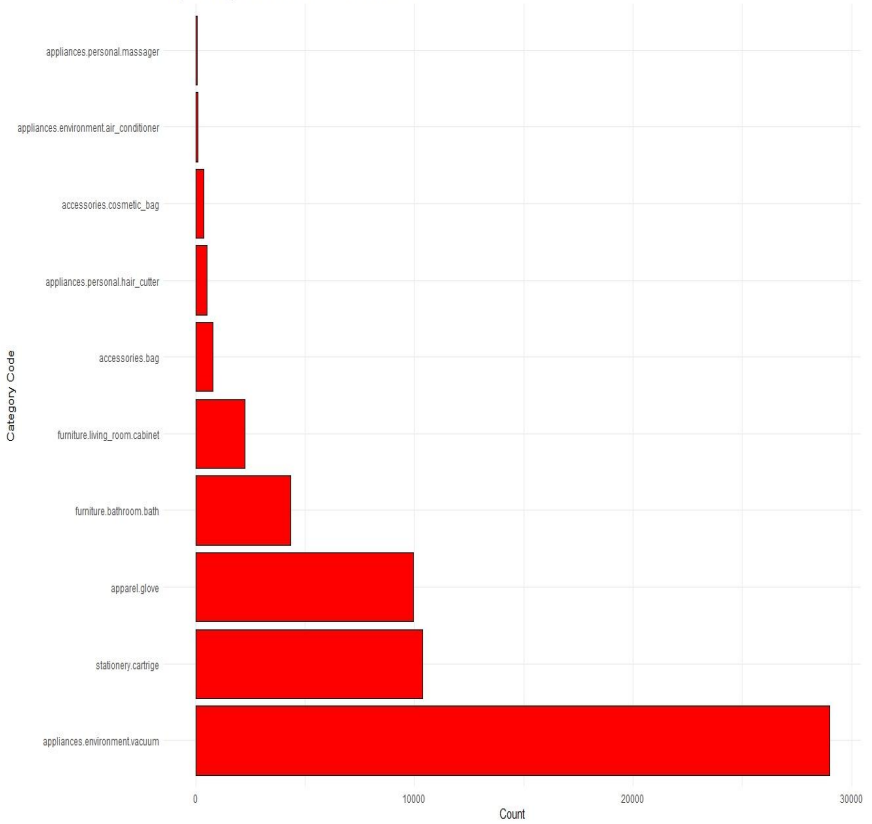
- **Analysis of user interaction patterns**
  E-commerce user interaction pattern analysis focuses on how customers navigate and engage with a website, identifying key behaviors like page visits, dwell time, and purchase pathways, which informs enhancements in website design, product placement, and personalized marketing strategies to better align with user preferences and address pain points.

- **Trends in product views and purchases**
  In User and Product Analysis through Exploratory Data Analysis, the examination of trends in product views and purchases reveals consumer interest and buying behavior patterns, aiding in demand forecasting, inventory optimization, and the customization of marketing strategies to align with customer preferences and trends.

- **Correlation between product prices and user behavior**
  Exploratory data analysis of user and product interactions focuses on the correlation between product prices and user behavior, shedding light on how pricing strategies affect consumer purchase decisions and aiding in the optimization of pricing models to boost sales and understand the price sensitivity of diverse customer segments.



Top 10 Brands in Terms of User interactions



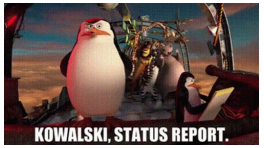
Top 10 Categories in Terms of User interactions

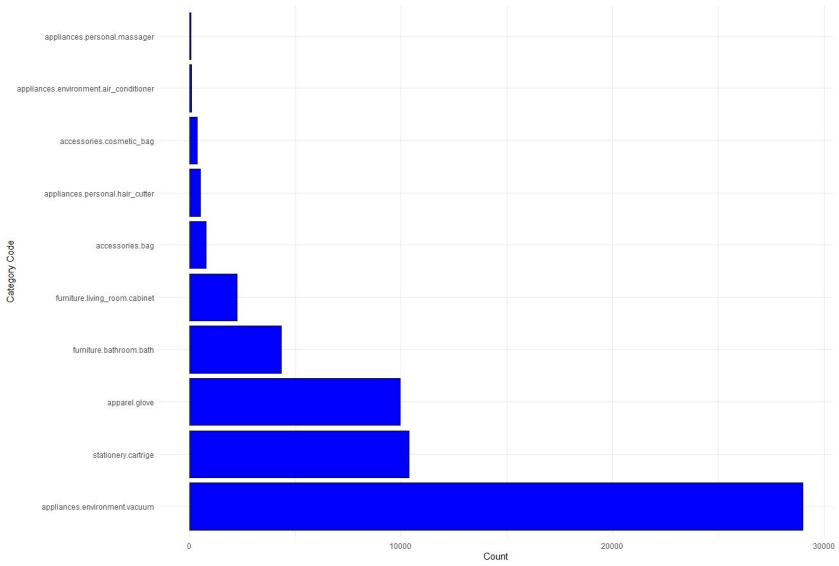
Exploratory Data Analysis - Category and Brand Analysis

- **Popularity and performance of different product categories**
  Exploratory data analysis in category and brand analysis involves evaluating the popularity and performance of various product categories through analysis of sales data, customer preferences, and market trends, providing critical insights for informed decision-making in inventory management,marketing strategies, and potential adjustments in product line-ups to meet consumer demand and market dynamics.

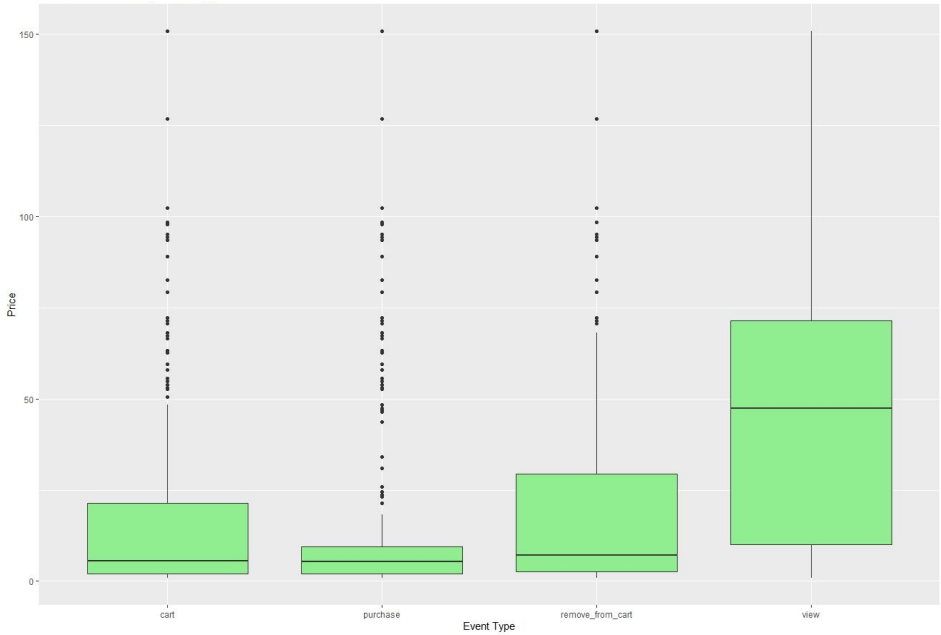
- **Brand influence on consumer choices**
  In Exploratory Data Analysis focusing on Category and Brand Analysis, the influence of brand on consumer choices is underscored by examining how brand reputation and recognition sway purchasing decisions, often outweighing factors like price and product features, thereby informing strategies to enhance brand loyalty and market positioning.

- **Cross-category brand comparison**
  Exploratory Data Analysis in Category and Brand Analysis often involves cross-category brand comparisons to understand how different brands perform across various product categories, revealing insights into brand strengths, market positioning, and opportunities for strategic brand extensions or adjustments.





Top 20 Product Categories



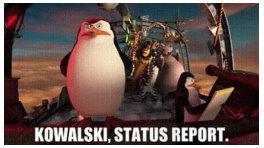
Box plot of prices by Event Type

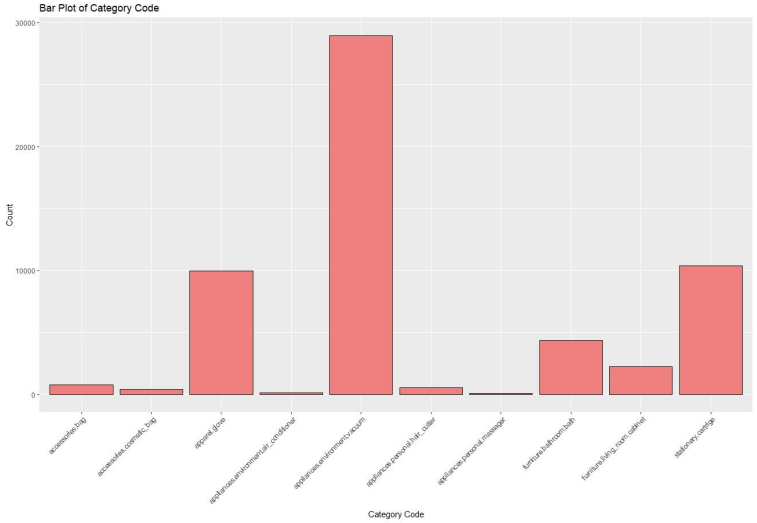
Exploratory Data Analysis - Customer Buying Habits

- **Analysis of purchasing patterns and trends**
  In exploratory data analysis focusing on customer buying habits, analyzing purchasing patterns and trends is crucial for identifying recurring behaviors, seasonal variations, and emerging trends in consumer buying, which helps businesses tailor their marketing and inventory strategies to meet evolving customer needs.

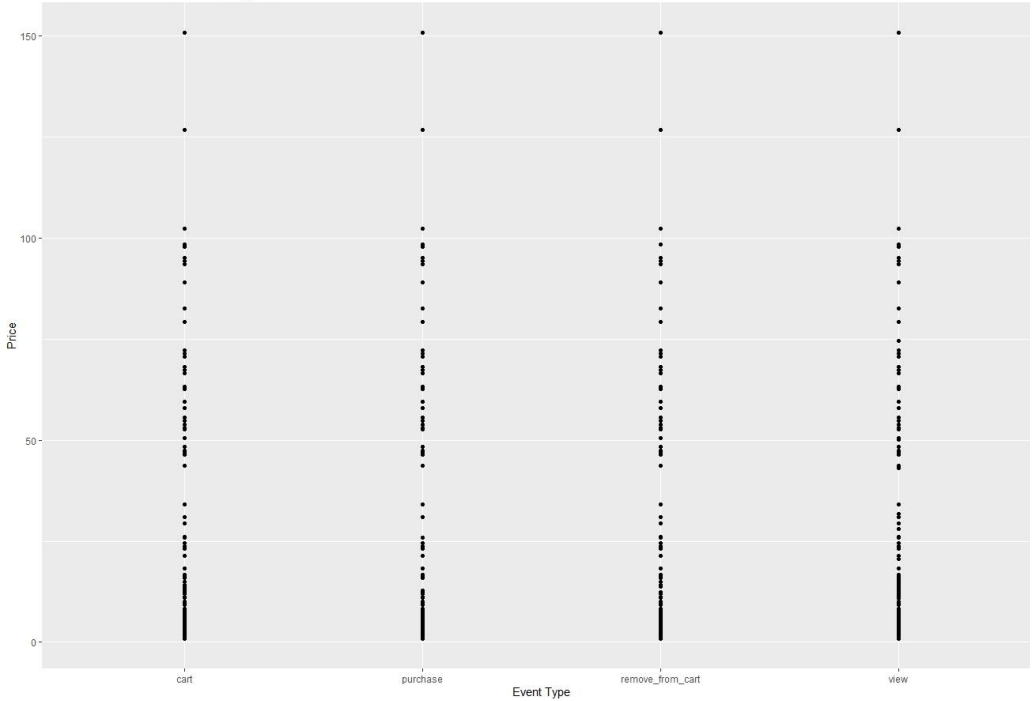
- **Impact of marketing and promotions on buying behavior**
  Under exploratory data analysis of customer buying habits, examining the impact of marketing and promotions on buying behavior is key to understanding how various advertising strategies and discount campaigns influence consumer purchase decisions, thereby guiding more effective and targeted marketing initiatives.

- **Customer loyalty and repeat purchase behavior**
  In exploratory data analysis focused on customer buying habits, investigating customer loyalty and repeat purchase behavior uncovers the underlying reasons for customers continued engagement and recurring purchases, essential for formulating strategies to enhance customer retention and loyalty.





Distribution of Customer-Product Interactions



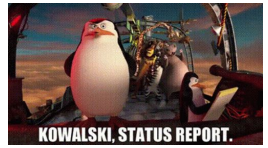
Scatter Plot of Price vs Event Type

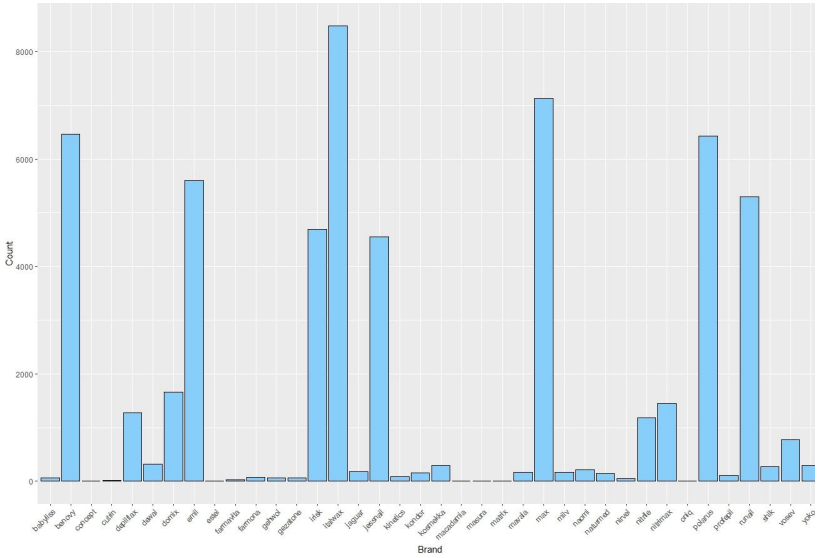
Exploratory Data Analysis - Brand Analysis

- ***Influence of brand recognition on customer decisions***
  In exploratory data analysis focusing on brand analysis, the influence of brand recognition on customer decisions is examined to understand how brand familiarity and reputation impact consumer purchasing choices, guiding marketing and branding strategies.

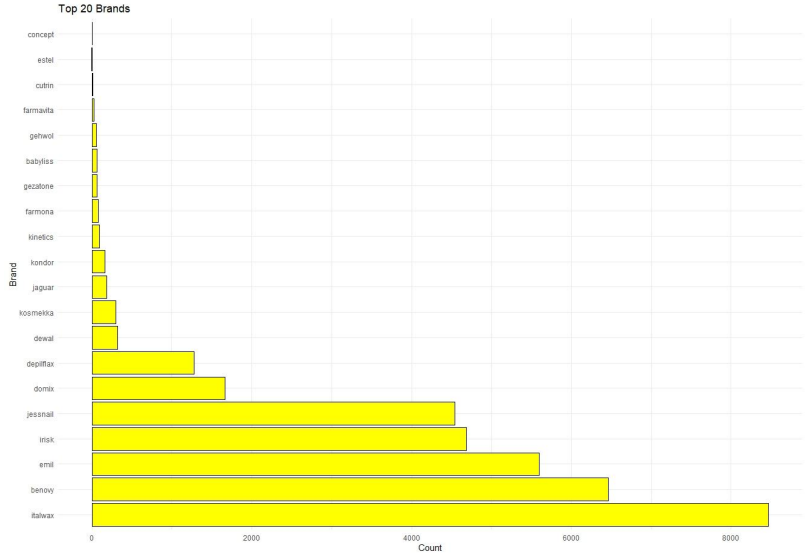
- ***Comparison of brand performance across different categories***
  Exploratory data analysis in brand analysis entails comparing the performance of Brand A across different categories to identify strengths and weaknesses in various market segments, informing targeted strategies for brand growth and positioning.

- ***Analysis of brand loyalty and customer engagement***
  The analysis of brand loyalty and customer engagement under exploratory data analysis in brand analysis highlights the relationship between a brand ability to foster loyal customers and their levels of engagement, providing insights for strategies to strengthen brand loyalty and enhance customer interaction.





Bar plot of Brand



Top 20 Brands

Feature Engineering Process

- Predictive Feature Identification - Utilizing statistical analysis and domain knowledge to select impactful predictors.
- Data Transformation Techniques - Implementing normalization, scaling, and one-hot encoding to standardize and categorize data.
- Feature Extraction - Applying Principal Component Analysis for dimensionality reduction and time-series decomposition for trend analysis.
- Interaction Features - Crafting combined features to explore complex inter-variable relationships.
- Handling Missing Data - Employing tailored imputation methods and algorithms adept at managing data gaps

Advanced Modeling and Statistical Techniques

- *Random Forest Prediction:* Implementation of Random Forest algorithm for both classification and regression tasks.
- *PCA (Principal Component Analysis)*: Utilized for dimensionality reduction and feature extraction.
- *Recursive Feature Elimination (RFE)*: Applied for selecting the most significant features for the models.
- *Statistical Tests*: Chi-square test for categorical variable independence, one-way ANOVA for comparing sample means, and independent samples t-test for mean comparison between two groups.

Hypothesis Testing and Baseline Model Development

- *Hypothesis Testing:* This involves using statistical methods to determine if there is enough evidence to support a specific theory or hypothesis about a dataset

- *Baseline Model Development:* The baseline model is a simple initial model used as a point of comparison for more complex models. Its performance, often measured in terms of accuracy or error rate, sets a benchmark for subsequent, more sophisticated models

Recursive Feature Elimination (RFE)

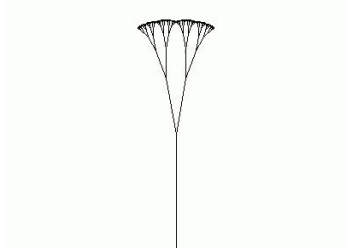
- *Feature Selection Process*: Utilization of RFE to identify the most impactful features for the models.

- *Model Efficiency Improvement*: By removing less important features, RFE helps in enhancing model performance and reducing complexity.

Statistical Tests for Data Insight

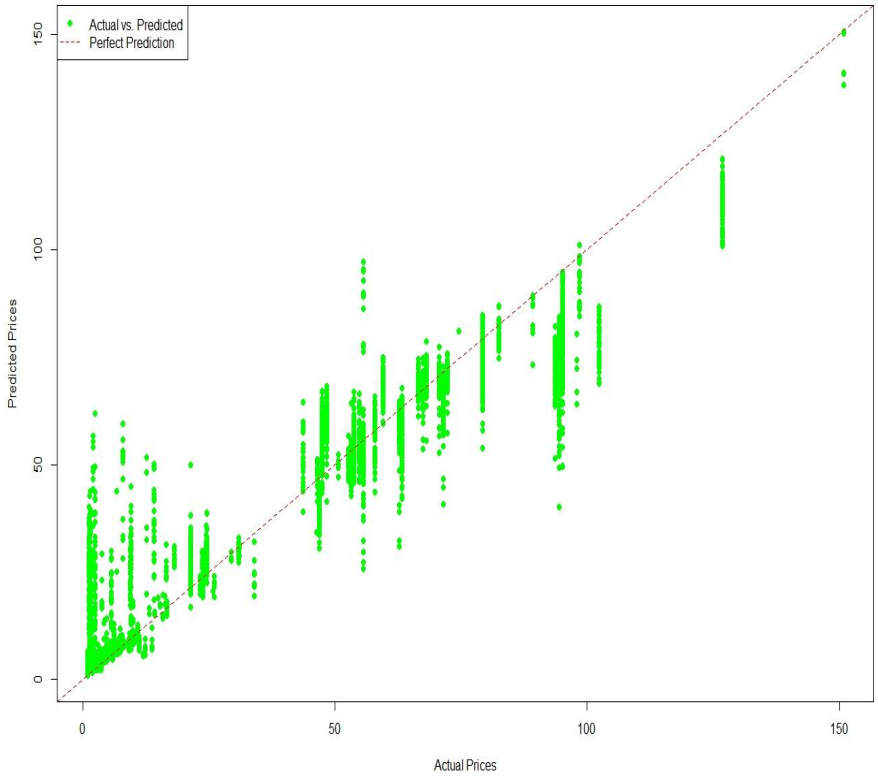
- *Chi-square Test for Independence:* Used to determine if there's a significant relationship between two categorical variables.
- *One-way ANOVA*: Applied to compare the means of three or more independent groups.
- *Independent Samples t-test:* Utilized to compare the means between two independent groups.

Random Forest and PCA Modeling

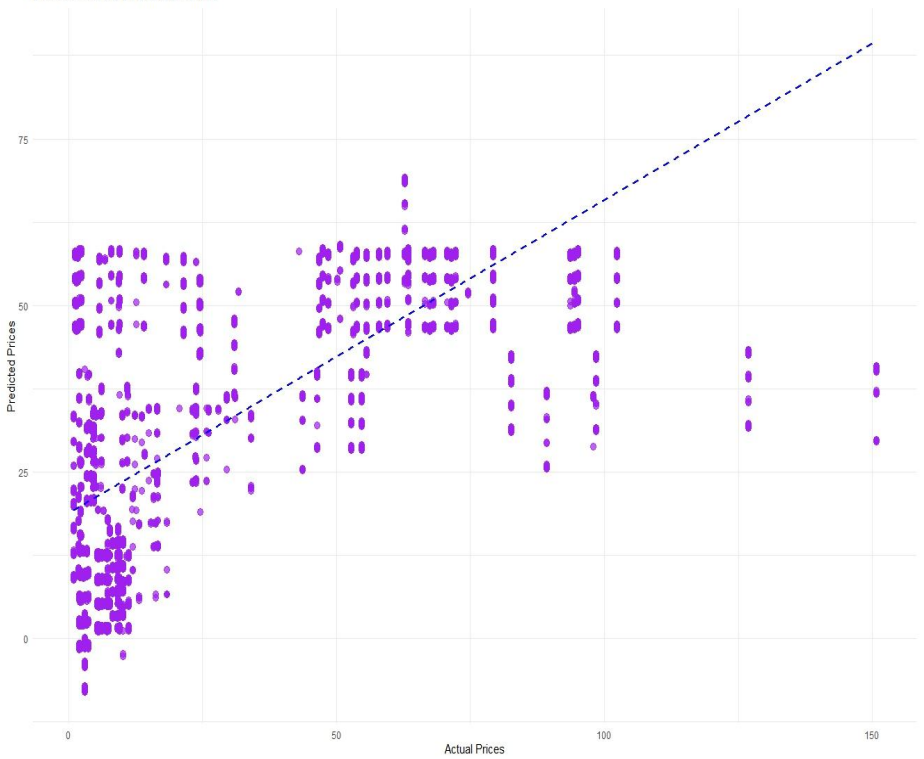


- *Random Forest Modeling:* Utilized for its robustness in handling both classification and regression tasks with high accuracy.

- *Principal Component Analysis (PCA):* Applied for reducing data dimensionality while retaining the most significant features for analysis.

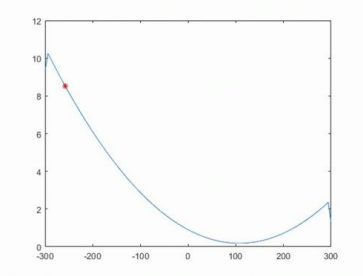


Random Forest : Actual vs Predicted Values



PCA: Actual vs Predicted Values.

Model Evaluation and Comparison



- *Evaluation Metrics:* Use of metrics like Mean Squared Error (MSE) and R-squared to assess model accuracy.

- *Model Comparison:* Analyzing the performance of different models to determine the most effective approach.

Future Work: Expanding Our Analytical Horizons

- *Advanced Analytical Techniques:* Exploration into AI and machine learning advancements for deeper insights.
- *Real-Time Data Analysis:* Implementing dynamic, real-time analytics for adaptive strategies.
- Longitudinal Studies: Conducting studies over extended periods to understand long-term trends and consumer behavior shifts.
- *Enhanced Personalization:* Leveraging data for more tailored user experiences and targeted marketing.
- *Scalability and Automation:* Focusing on scaling analytical models and automating processes for efficiency.

Conclusion

- *Customer Insights: Through diverse visualizations, we uncovered concentrated customer interactions, emphasizing specific brands and affordable items.*
- *Behavioral Trends: Analysis revealed browsing behavior dominance, implying potential areas for marketing strategy enhancement.*
- *Pricing Dynamics: The price distribution showcased a price-sensitive customer base, aiding in the formulation of targeted pricing strategies.*
- *Predictive Modeling: Leveraging Random Forest and PCA, we explored the accuracy of price predictions, providing valuable insights into model performance.*

.

Thank You