

A photograph showing a long, diagonal line of red metal shopping carts parked on a concrete surface. The carts are arranged in a neat row, receding into the distance. The perspective is from a low angle, looking down the line of carts. The carts have a standard design with a wire mesh basket and four wheels. A semi-transparent dark grey banner is overlaid on the lower left portion of the image, containing the title and author's name in white and black text.

Consumer Behavior and Purchasing Intentions in E-Commerce

OMONIYI OLUWATEMI LORUN MATTHEW

Data Description

- **Source:** UCI Machine Learning Repository
- **Dataset:** "Online Shoppers Purchasing Intention Dataset"
- **Observations:** 12,330 rows
- **Features:** 17 columns
- **Target Variable:** Revenue (*Boolean*)
- **Features of Interest:** BounceRates, PageValues, ProductRelated, ExitRates, etc.
- **Focus:** Shopper browsing behavior and purchasing intention

Features	Data Types
<ul style="list-style-type: none">• Administrative• Informational• ProductRelated• OperatingSystems (Categorical)• Browser, Traffic Type, Region (Categorical)	INTEGER
<ul style="list-style-type: none">• Administrative_Duration• Informational_Duration• ProductRelated_Duration• BounceRates, ExitRates• PageValues• SpecialDay	FLOAT
<ul style="list-style-type: none">• Revenue• Weekend	BOOLEAN
<ul style="list-style-type: none">• VisitorType	OBJECT



Research Questions & Motivation

Question 1 (Supervised)

- **QUESTION:** *"What factors influence a shopper's likelihood to make a purchase?"*
- **MOTIVATION:** *Improve marketing strategies and user experience to boost conversion rates.*

Question 2 (Unsupervised)

- **QUESTION:** *"Can we identify distinct groups of shoppers based on their browsing behavior?"*
- **MOTIVATION:** *Enable customer segmentation for personalized recommendations and better resource allocation.*



Data Preparation

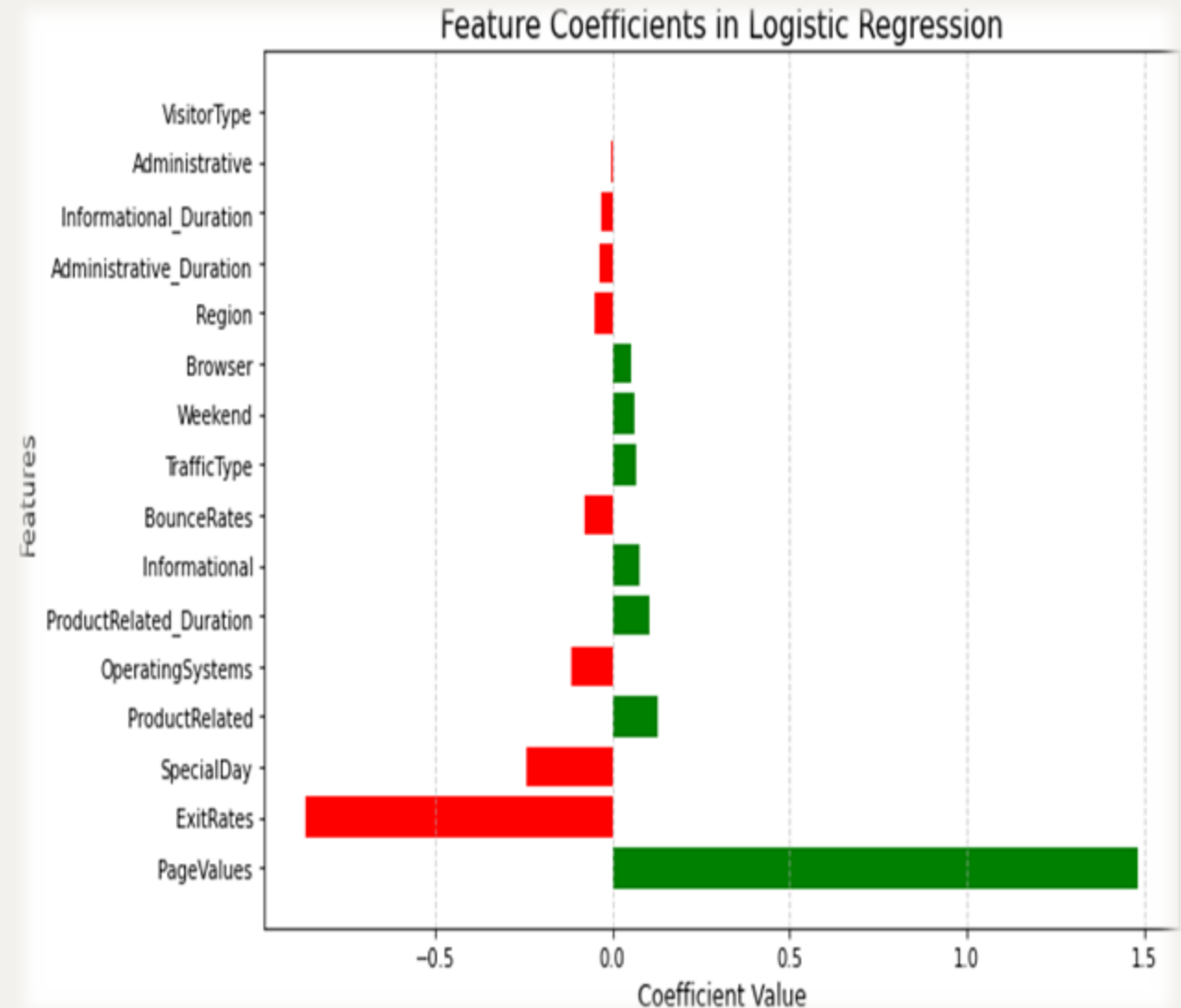
- **MS Excel Spellcheck**
- **Verified** no missing values for clean analysis.
- **Transformed** "Weekend" and "Revenue" columns to Boolean for easier processing
- **Standardized** feature scaling for clustering and logistic regression models

Supervised Model Results – Logistic Regression

- **Top Positive Predictors:** *PageValues* (+1.48), *ProductRelated* (+0.13)
- **Top Negative Predictors:** *ExitRates* (-0.86), *SpecialDay* (-0.24)
- **Model Accuracy:** 86.98%

Insights:

- High **PageValues** indicate pages contributing to revenue positively
- High **ExitRates** lower the likelihood of purchase

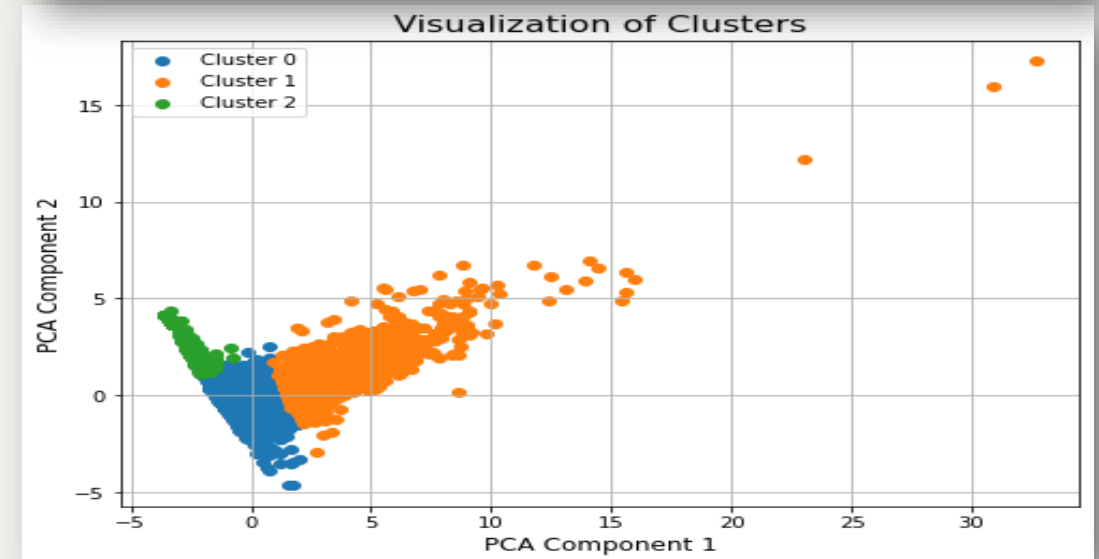
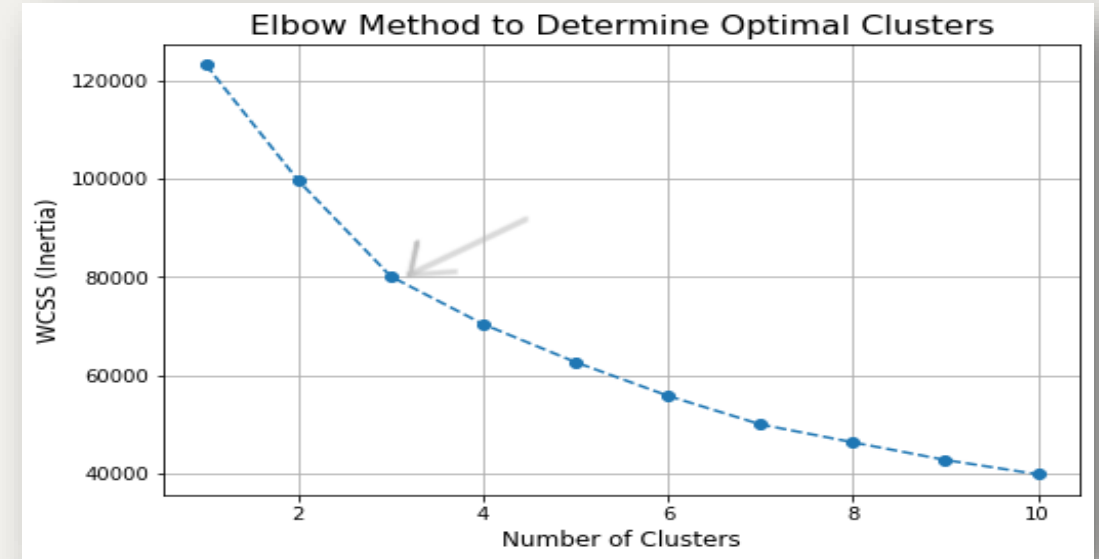


Unsupervised Model Results - Clustering

- Optimal clusters were determined using the **Elbow Method**: 3 clusters.

CLUSTER SUMMARIES

- **CLUSTER 0**: Low engagement across features, **unlikely purchasers**
- **CLUSTER 1**: High ProductRelated interaction and PageValues, **likely purchasers**
- **CLUSTER 2**: Moderate engagement, **potential purchasers with targeted effort**
- Clusters could provide actionable insights for tailoring marketing strategies.





Conclusions and Insights

- **Supervised Learning:** Key factors like PageValues and ExitRates significantly influence purchasing decisions
- **Unsupervised Learning:** Shopper segmentation reveals distinct groups of shoppers based on their browsing behaviors
- Insights empower businesses to optimize marketing, enhance website design, and increase ROI
- Highlighted the importance of personalized customer experiences for improving satisfaction

THANK YOU!
DO YOU HAVE ANY QUESTIONS?

