# **Logistic Regression Analysis**

**PART A: NON-TECHNICAL REPORT**

**OBJECTIVE**

The aim of this research is to predict how many of Fresco's loyalty cardholders can be categorized as low spenders (£50 or less per week) and high spenders (£50+). It will help the marketing team identify customer segments with which they should implement targeted promotional strategies.

**DATA SUMMARY**

The data set contains weekly records of shopping behaviors of customers for a period of 26 weeks, including demographic details, store types, shopping frequency, and product preferences (value, brand, and premium Fresco Top).  
A binary variable was created to classify expenditure, with 0 indicating low spenders and 1 indicating high spenders.

**METHODOLOGY**

* Binary logistic regression was utilized as an appropriate methodology for classifying customers into clear groups of low and high spending.
* Elimination procedure was applied through which all predictors without statistical significance were eliminated, and those with considerable predictive power remained:  
  • Value-added Products  
  • Commercial goods

**KEY FINDINGS**

The model has high predictive accuracy (94.7%).

* Customers purchasing more brand and value products are more likely to be high spenders.
* The chance of becoming a major consumer increases by 1.9 times for each additional brand product and by 1.6 factor for every supplemental value product.

**RECOMMENDATIONS**

* Promote value and brand products to increase average basket value.
* Identify and target potential high-value customers through customized loyalty rewards.
* Implement the model on digital media channels to enable real-time targeting.

**PART B: TECHNICAL ANALYSIS REPORT**

1. **Data Preparation**

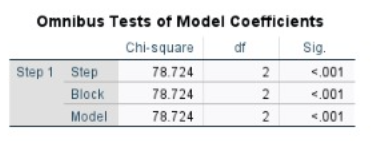
* Imported and cleaned dataset from Fresco Supermarket.
* Created a new binary variable: HighSpender (0 = ≤£50, 1 = >£50).
* Carrying out logistic regression using the Enter approach and then removing non-statistically significant variables.

**2. Model Selection & Assumptions**

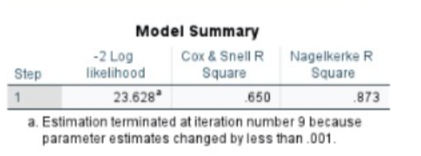
* Binary logistic regression is suitable for dichotomous outcomes.
* Assumptions:
  + **No multicollinearity**: Insignificant predictors were removed.
  + **Linearity in log odds**: Not a concern for categorical/predictor counts.
  + **Adequate sample size**: >40 records ensures stable estimates.

**3. Final Model Output**

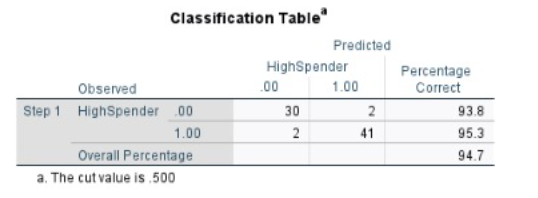
* **Omnibus Test of Model Coefficients**: χ² = 78.724, p < 0.001



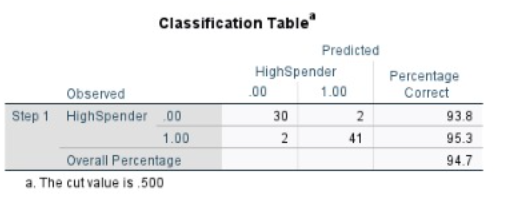
* **Model Summary**:
  + Cox & Snell R² = 0.650
  + Nagelkerke R² = 0.873



* **Classification Accuracy**:
  + Overall: 94.7%
  + Low Spenders: 93.8%
  + High Spenders: 95.3%



**4. Coefficient Interpretation**

****

* **Value Products**
  + B = 0.460, p = 0.004
  + Exp(B) = 1.585 → Odds of being a high spender increase 1.6x per unit increase.
* **Brand Products**
  + B = 0.641, p = 0.003
  + Exp(B) = 1.900 → Odds nearly double per unit increase.

**CONCLUSION:**

The logistic regression model provides an extremely accurate method for the classification of consumer expenditure patterns. Through the use of only two important variables, the model keeps a simplicity that increases its effectiveness, making it ideal for implementation within Fresco's campaign management and customer segmentation systems.