



Improving overall sales performance:

Python and R Analysis

Presented by Eloise Farmer

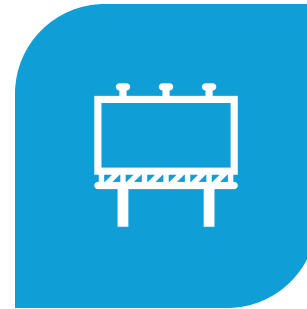
Focus Questions:



HOW CUSTOMERS EARN AND
USE LOYALTY POINTS



HOW WE CAN SEGMENT
CUSTOMERS INTO CLEAR
GROUPS FOR TARGETED
MARKETING

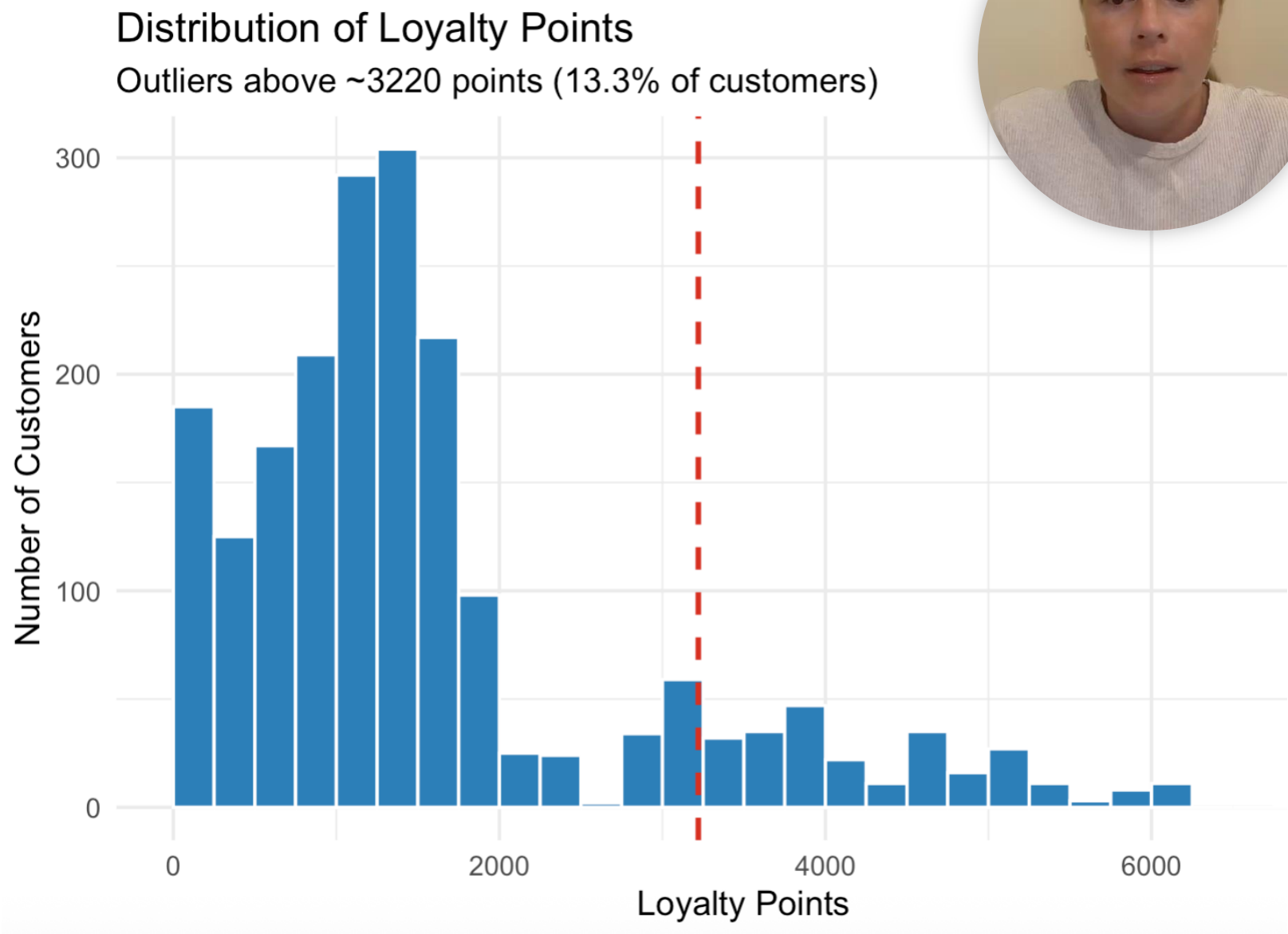
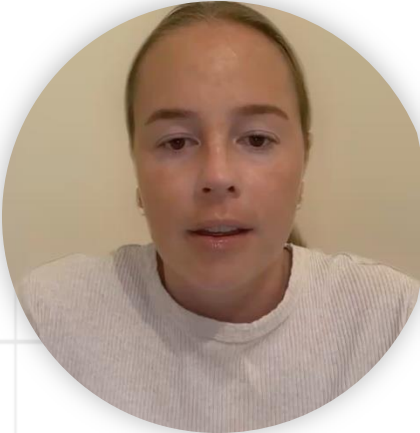


HOW REVIEWS CAN INFORM
CAMPAIGNS AND PRODUCT
IMPROVEMENTS



IS LOYALTY DATA IS SUITABLE
FOR PREDICTIVE MODELLING

Exploratory Data Analysis:



Total Customers: 2000

Key Drivers of Loyalty Points Accumulation:

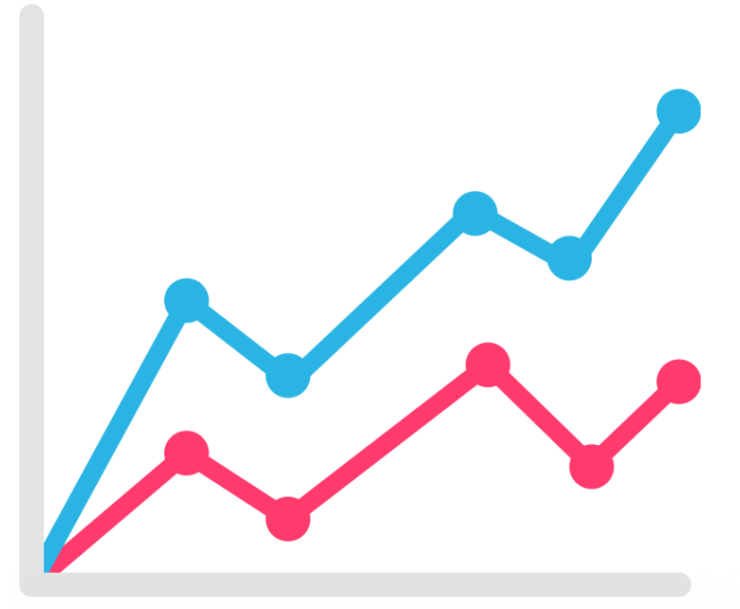
Income and Spending score



Correlation Matrix of Key Variables



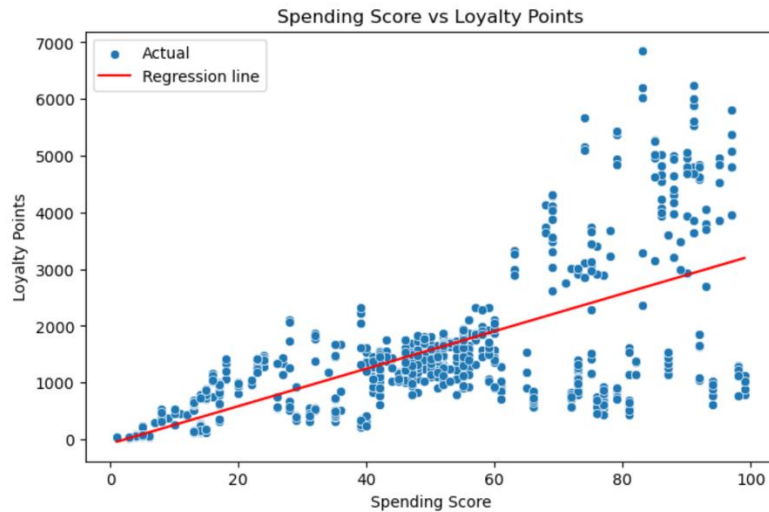
Predictive Modelling: From Description to Prediction



Simple Linear Regression Modelling

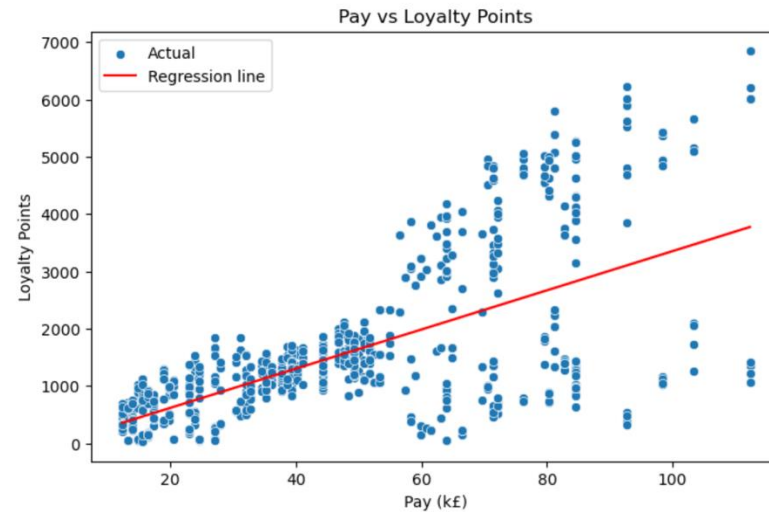


Spending



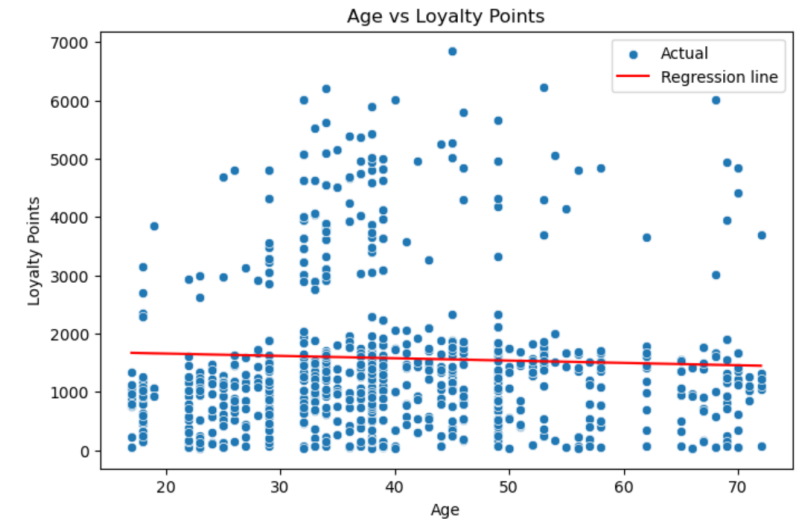
$$R^2 = 45\%$$

Pay



$$R^2 = 38\%$$

Age

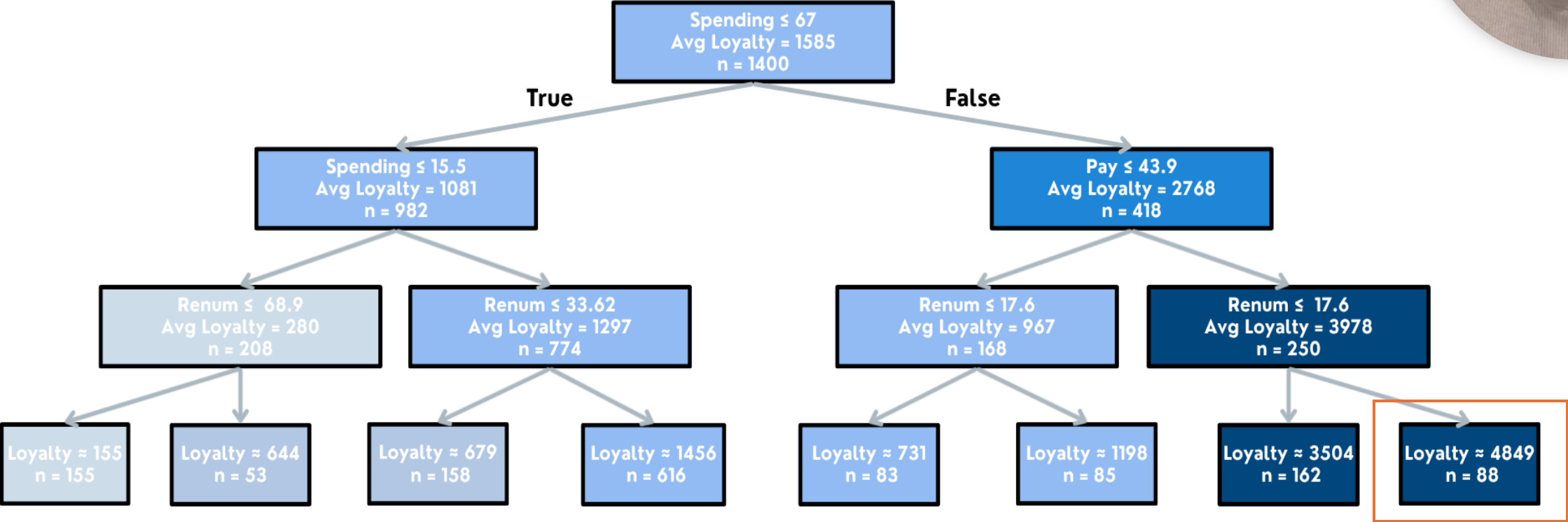


$$R^2 = 0.2\%$$

RMSE \approx 1000

MSE \approx 700

Decision Tree



Depth 3
 $R^2 = 92\%$
RMSE = 371
MAE = 267

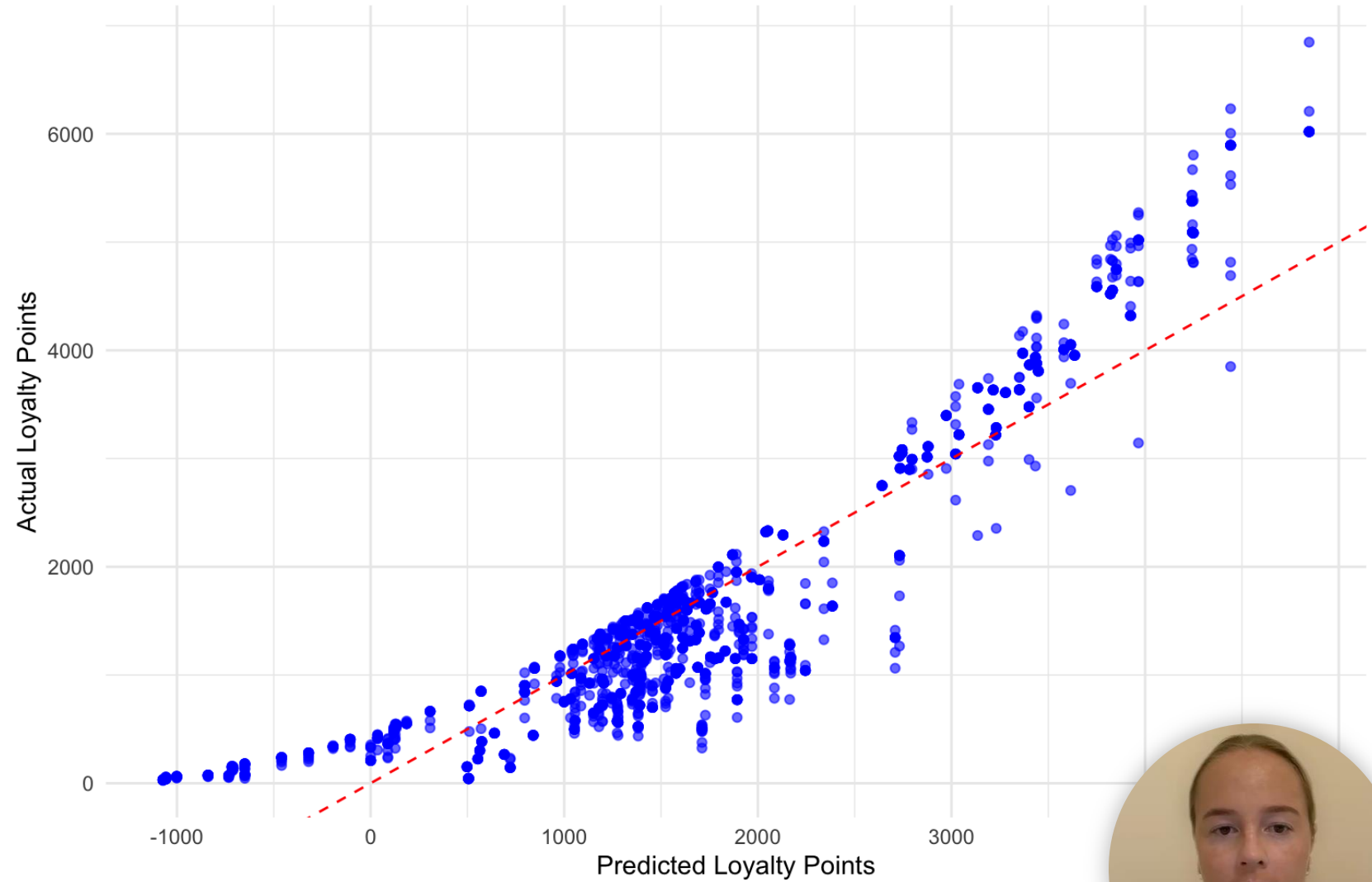
Multiple Linear Regression

$R^2 = 83\%$

RMSE = 530

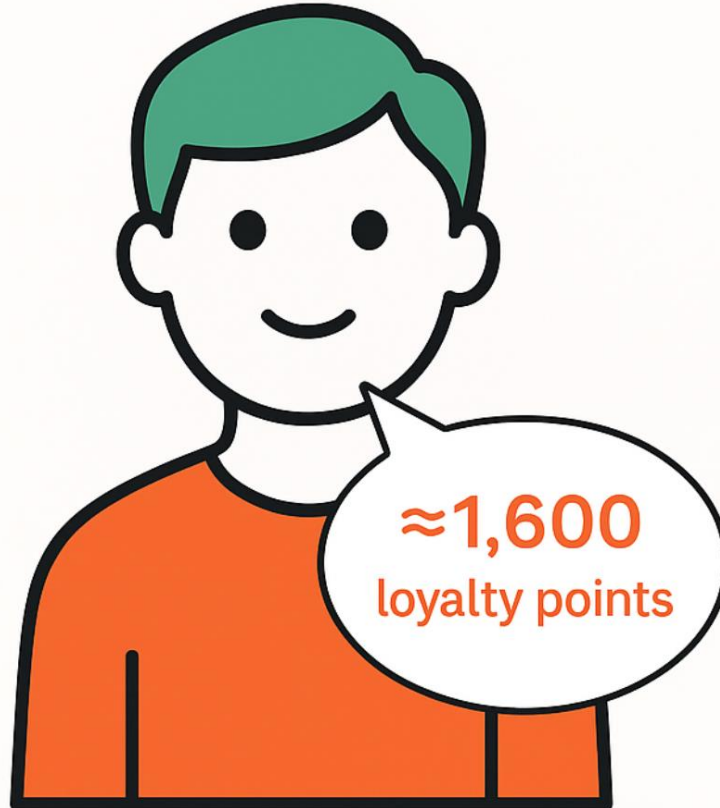
MAE = 412

Multiple Linear Regression: Actual vs Predicted Loyalty Points



MLR Example:

Pay £45,000
Spending score 55



- Segment customer early
- Target campaigns
- Increase Sales



Key Customer Clusters Overview

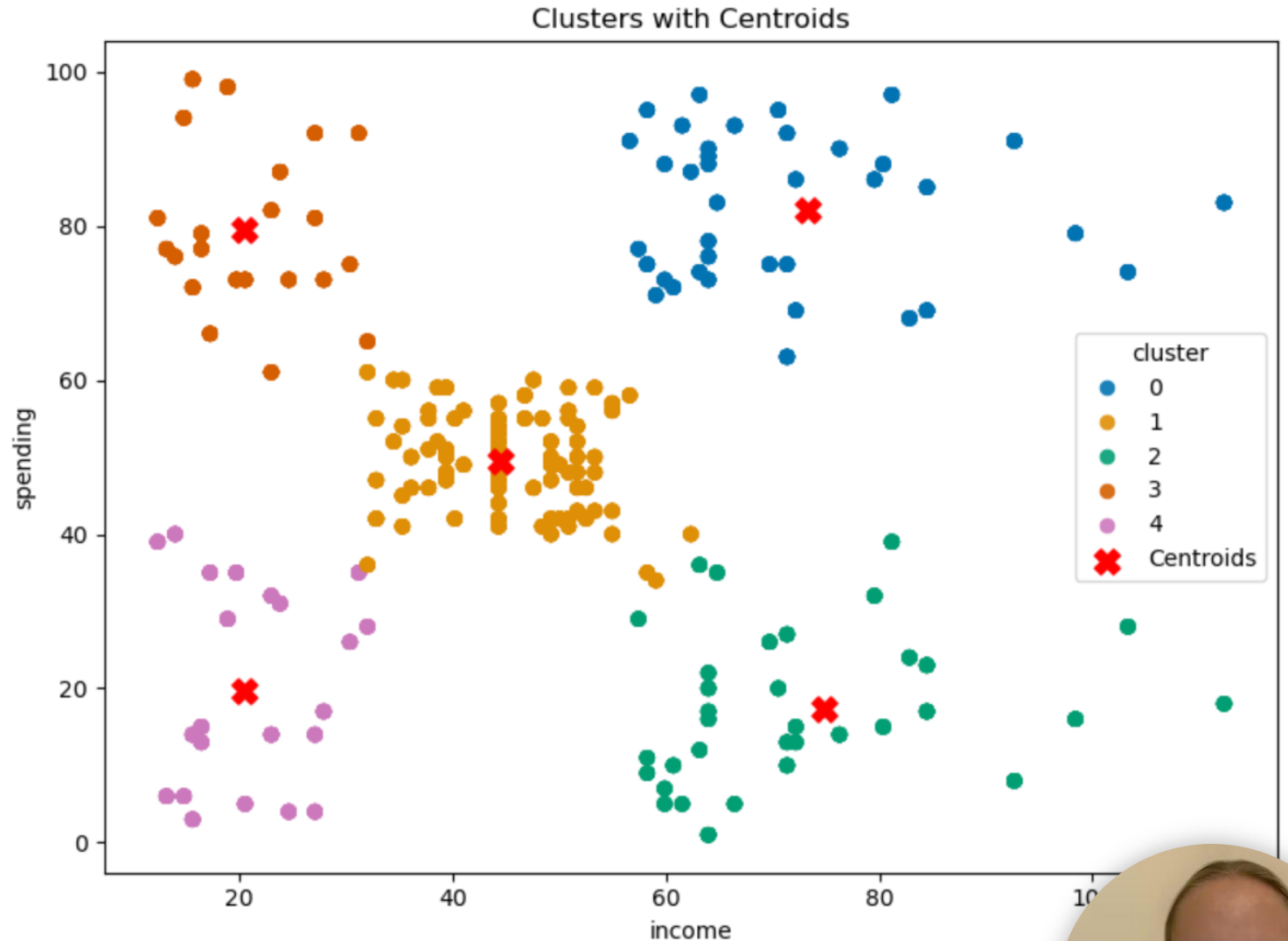
Cluster 0 High income, high spend

Cluster 1 Middle income, average spend

Cluster 2 High income, low spend

Cluster 3 Low income, high spend

Cluster 4 Low income, low spend

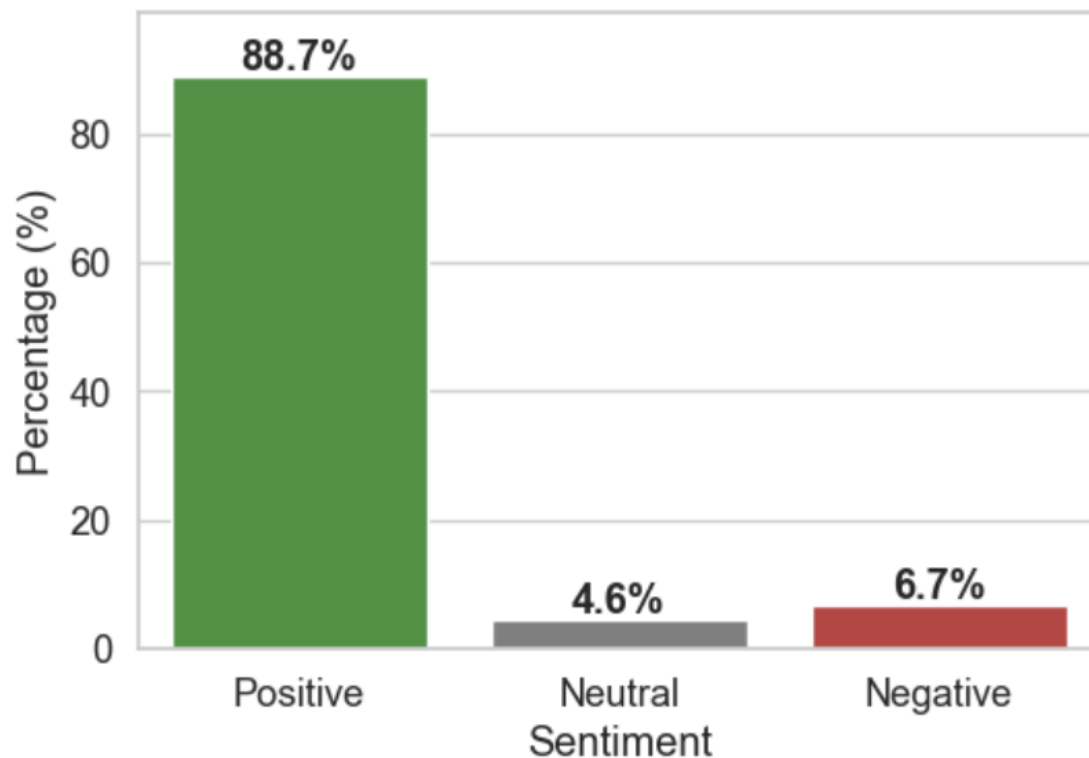


Sentiment Analysis:

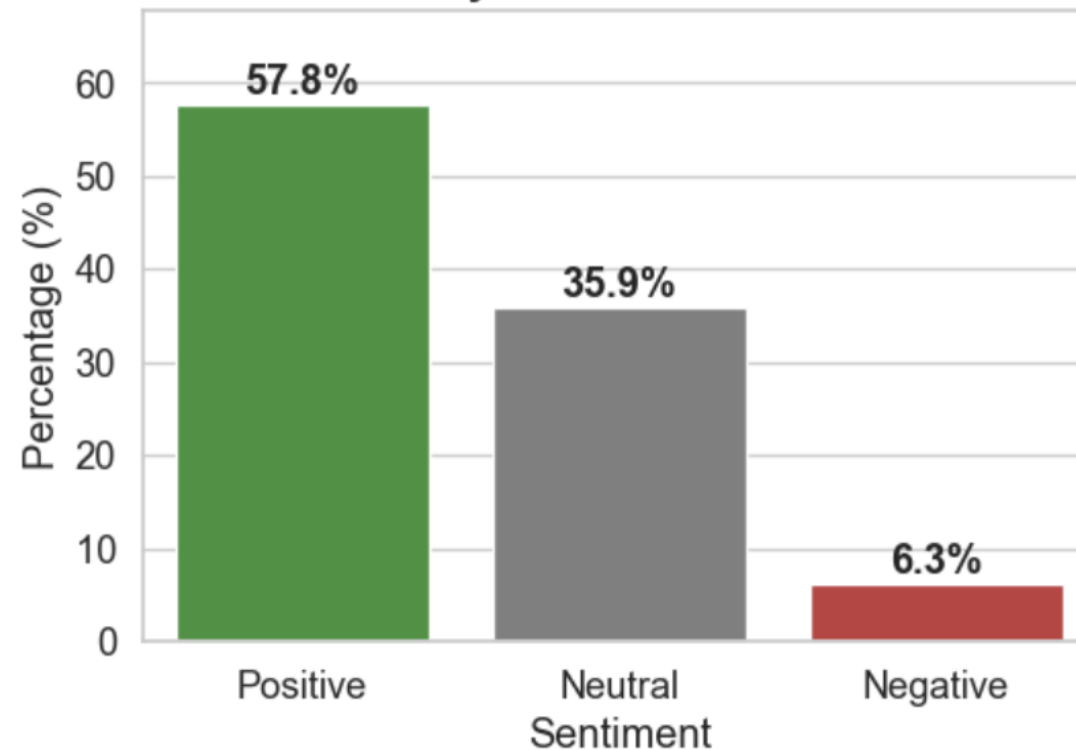
Customer Satisfaction Score (CSAT): 73%

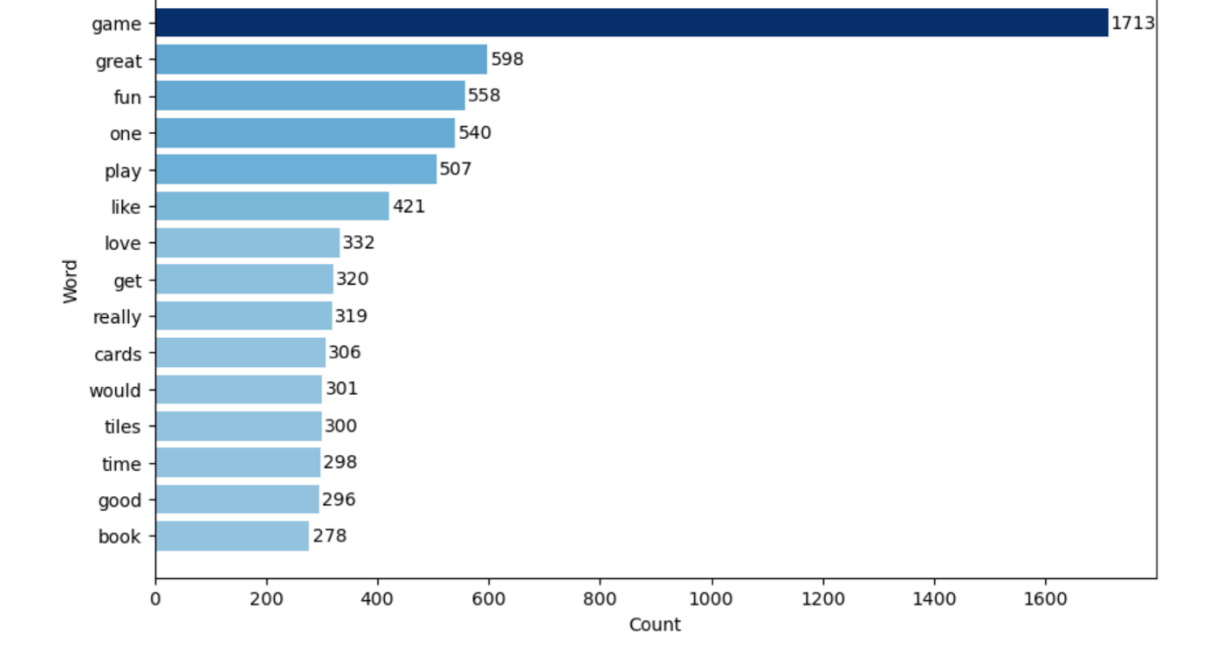
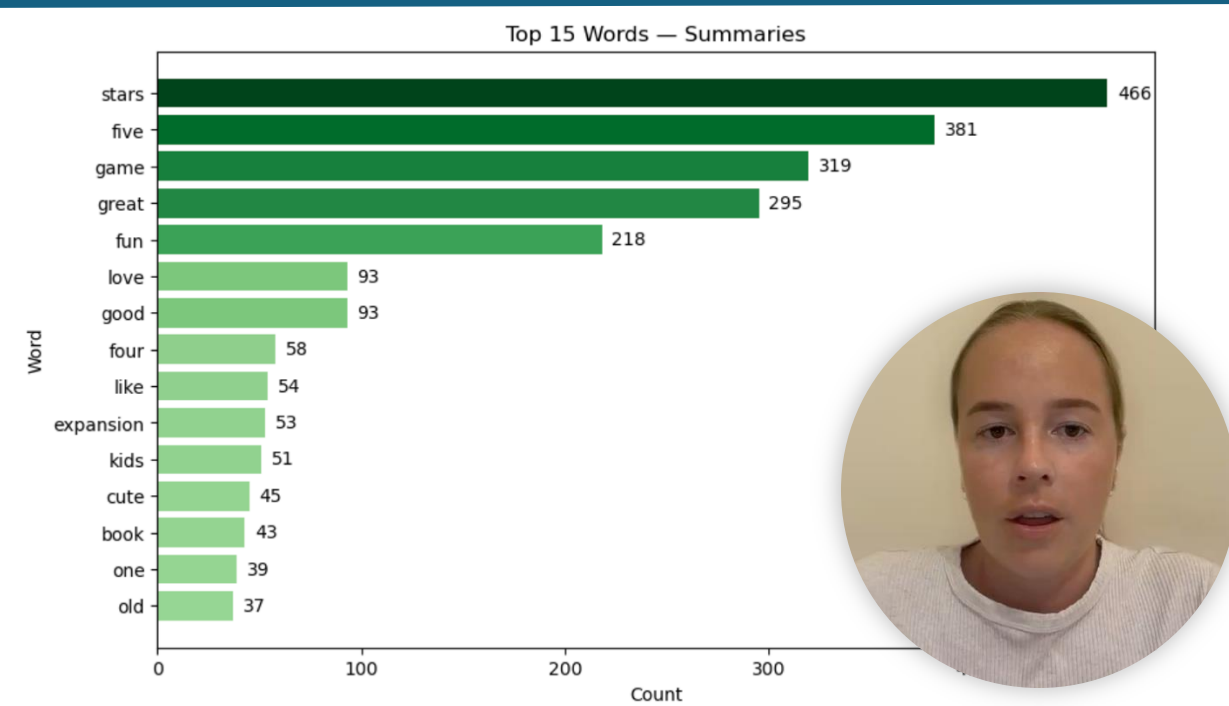


Review Sentiment Distribution



Summary Sentiment Distribution



[illegible][illegible]

Product Analysis

- Spending Score 1-100 as a proxy
- VADER Sentiment Scores



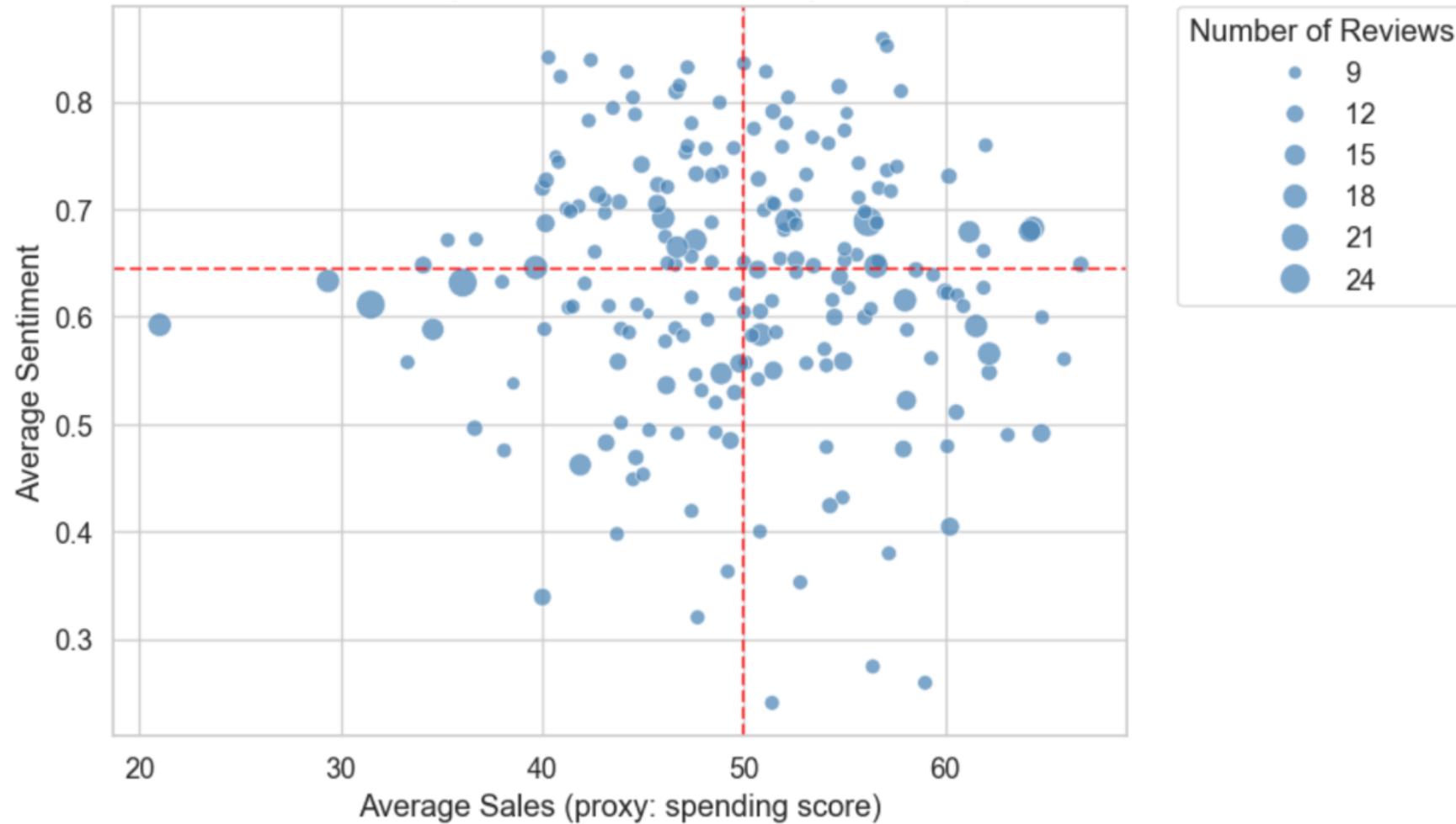
Product Analysis:

Top & Bottom Performers



Category	Products (scores)
Top 3 by Sales	3153 (~67) 5510 (~66) 6466 (~65)
Bottom 3 by Sales	1031 (~31) 1175 (~29) 2173 (~21)
Top 3 by Sentiment	2874 (~0.86) 5429 (~0.85) 5493 (~0.84)
Bottom 3 by Sentiment	2253 (~0.27) 3165 (~0.26) 9597 (~0.24)

Product Analysis: Sales vs Sentiment



Quadrant Interpretation: Sales vs Sentiment

Quadrant	Product Code (sales, sentiment)		Business Action
💰 Cash Cows (High Sales, High Sentiment)	3153	(66.7, 0.65)	Maintain investment, loyalty programs, ensure supply chain reliability
	4405	(64.4, 0.68)	
	5726	(64.2, 0.68)	
🌟 Hidden Gems (Low Sales, High Sentiment)	5493	(40.3, 0.84)	Increase promotion and distribution to scale successes
	2371	(42.4, 0.84)	
	3427	(47.2, 0.83)	
⚠️ Risks (High Sales, Low Sentiment)	5510	(65.9, 0.56)	Investigate quality issues, address complaints, improve perception
	6466	(64.8, 0.60)	
	2829	(64.8, 0.49)	
❌ Low Priority (Low Sales, Low Sentiment)	2173	(21.0, 0.59)	Consider rationalisation, low-cost engagement, or discontinuation
	1175	(29.4, 0.63)	
	1031	(31.5, 0.61)	



Recommendations

Deploy Multiple Linear Regression

- Accurate, interpretable, predicts expected loyalty



Segment Customers Immediately

- Avoid one-size-fits-all campaigns



Segment Actions

- Premium: Retain with VIP rewards & exclusivity
- Affluent but disengaged: Re-engage with personalised offers
- Mainstream: Seasonal campaigns & referrals
- Value seekers: Bundles & multipliers, but market responsibly
- Budget-conscious: Low-cost automated engagement



Recommendations

Product-Level Strategy

- 💰 Protect Cash Cows (loyalty perks, supply assurance)
- 🌟 Promote Hidden Gems (campaigns, cross-sell)
- ⚠️ Investigate Risk Products (high sales, poor sentiment)
- ❌ De-prioritise Low Priority products
 - Investigate worst-reviewed product codes

Customer Feedback

- Phase out short summaries → replace with star ratings
- Continue reviews for rich, qualitative insight
- Real-time NLP monitoring to catch issues early



Further Analysis



EXPAND DATA
COLLECTION:
PURCHASE HISTORY,
PREFERENCES,
CROSS-PLATFORM



LINK LOYALTY &
SENTIMENT DIRECTLY
TO SALES REVENUE
AND PRODUCT NAMES



ANALYSE SEASONAL &
REGIONAL PATTERNS
IN LOYALTY AND
SENTIMENT



CONNECT LOYALTY
POINTS TO CUSTOMER
LIFETIME VALUE



EXPLORE ADVANCED
MODELS (E.G.,
RANDOM FORESTS)
FOR HIGHER
ACCURACY

