



Optimizing Grocery Marketing Campaigns

By: Elaina Manalac &
Jacob Manzella

Table of contents

01

Introduction

Business problem & stakeholders

02

Data Collection

Data cleaning, transforming, and construction of new variables

03

Analysis & Results

Use of statistical methods, linear regression, & visualizations

04

Conclusion

Implications & future extensions





01

Introduction

Business problem & stakeholders

Problem Statement

This grocery store **struggles** to make their **promotional strategies effective**, leading to **wasted marketing budgets** and **low response rates**. Without the ability to **predict which customers** are likely to respond to offers, this store risks spending money on **promotions that don't work**. This inefficiency affects **profits** and **competitiveness**, highlighting the need for a **better way to target customers** and **improve marketing results**.



Importance of the Problem



Financial Impact

Substantial financial losses from ineffective promotions and wasted spending.



Market Competitiveness

Gaining a competitive edge through targeted promotions for increased market share and loyalty.



Resource Allocation

Improved budget allocation for high-value marketing initiatives.



Stakeholders



Marketing & Sales Team

Use data to understand customer segments, evaluate campaigns, and refine targeting for branches..



Data Analyst Team

Provide insights and recs to other departments to inform strategic decisions.



CRM Team

Optimize customer interactions and address complaints to improve satisfaction.



Finance Team

Analyze income and purchasing behavior to evaluate profitability across customer segments.



Product Development

Tailor product offerings based on purchasing behaviors.



Executive Leadership

Use summary insights to guide business strategy and growth.





02

Data Collection

Data cleaning, transforming, and construction of new variables





2270

Rows

27

Columns

16

Numeric

income, wine, meat,
recency, web_purchases,
store_purchases...

3

Categorical

education, marital_status,
customer_enrollment

8

Dummy

kids, teen,
accepted_1...5

Data Dictionary

accepted_1 through accepted_5 - 1 if customer accepted the offer in the 1st, 2nd, 3rd, 4th, or 5th campaigns, 0 otherwise

response (target) - 1 if customer accepted the offer in the last campaign, 0 otherwise

complain - 1 if customer complained in the last 2 years

customer_enrollment - date of customer's enrollment with the company

education - customer's level of education

marital_status - customer's marital status

kids - number of small children in customer's household

teen - number of teenagers in customer's household

income - customer's yearly household income

fish - amount spent on fish products in the last 2 years

meat - amount spent on meat products in the last 2 years

fruits - amount spent on fruits products in the last 2 years

sweet - amount spent on sweet products in the last 2 years

wine - amount spent on wine products in the last 2 years

gold - amount spent on gold products in the last 2 years

deals_purchases - number of purchases made with discount

catalog_purchases - number of purchases made using catalogue

store_purchases - number of purchases made directly in stores

web_purchases - number of purchases made through company's web site

web_visits - number of visits to company's web site in the last month

recency - number of days since the last purchase



Limitations

Location & Size

The regional market and company scale can significantly influence customer preferences, purchasing behavior, and profitability.

Marketing Channels

Delivery method/ channel of campaigns was unknown.

Campaign Details

The specific focus and content of the five campaigns are not provided, making it difficult to assess their impact.

Economic Factors

Market conditions and seasonal trends can have a notable impact on purchasing behavior.



A vibrant, diagonal slice of various vegetables serves as the background for the top-left portion of the slide. It includes white cauliflower florets, a large orange tomato, a red bell pepper, and some green herbs like dill and cilantro.

03

Analysis

Use of statistical methods, linear regression, & visualizations

Correlation Matrix

Began with a correlation matrix to get a sense of relationships within our data

Top Positive Relationships:

Meat and catalog_purchases

Wine and store_purchases

Wine and catalog_purchases

Top Negative Relationships:

Income and website_visits

Meat and website visits

Kid_home and catalog_purchases

Top 10 Strongest Positive Relationships:

Variable 1	Variable 2	Correlation
meat	catalog_purchases	0.7238265908738766
wine	store_purchases	0.6421004382014517
wine	catalog_purchases	0.6352257380760022
fruits	fish	0.5948035099014977
income	catalog_purchases	0.5891624419343273
income	meat	0.5846333567663218
fish	sweet	0.5798700810242032
income	wine	0.5786497501367387
meat	fish	0.5684021297555203
fruits	sweet	0.5671639397753481

Top 10 Strongest Negative Relationships:

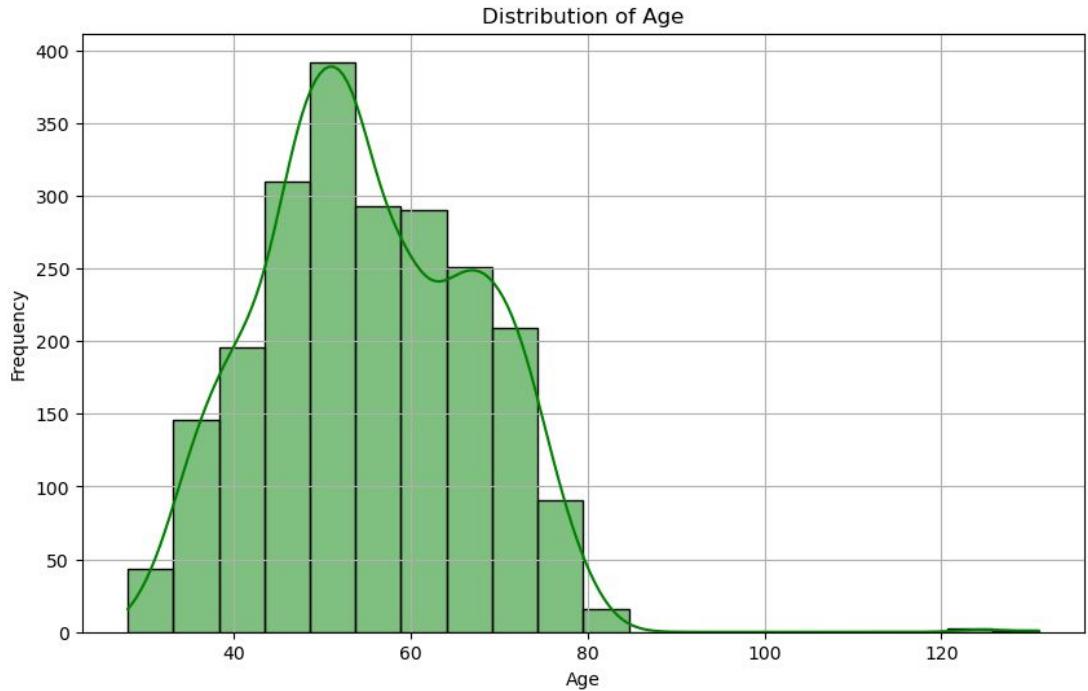
Variable 1	Variable 2	Correlation
income	website_visits	-0.5530880116530972
meat	website_visits	-0.5394698236237003
catalog_purchases	website_visits	-0.5203637147218997
kid_home	catalog_purchases	-0.5022372229807834
kid_home	store_purchases	-0.4996825649958032
kid_home	wine	-0.49629692675798376
fish	website_visits	-0.44600265439673764
kid_home	meat	-0.43712948235475135
income	kid_home	-0.42866900796918467
store_purchases	website_visits	-0.42847252671890473

Age Distribution

Histogram was an easy initial visualization to get a sense of the demographic of our data.

Age is normally distributed variable.

Most customers fall within the 50-55 age range with the youngest between 30- 35 and oldest between 80-85 years old.

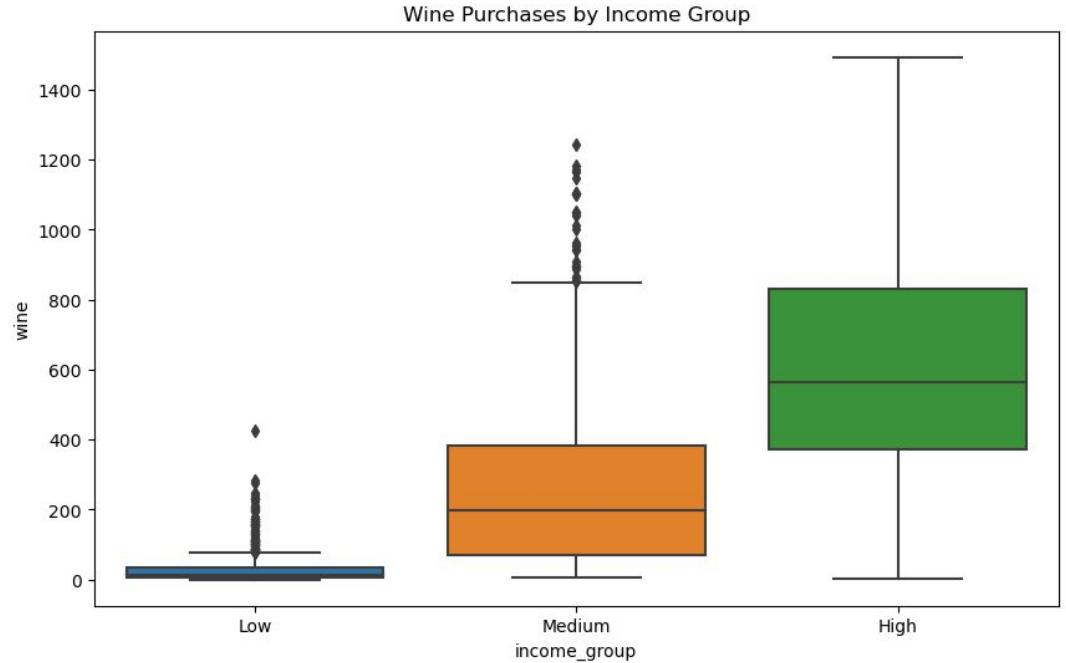


Customer Segmentation Based on Income

Box plot shows the distribution of wine purchases across 3 income groups.

Created income_group variable based on quantiles.

High Income Group: largest range of wine purchases and the highest median indicating customers with higher income purchase more wine.

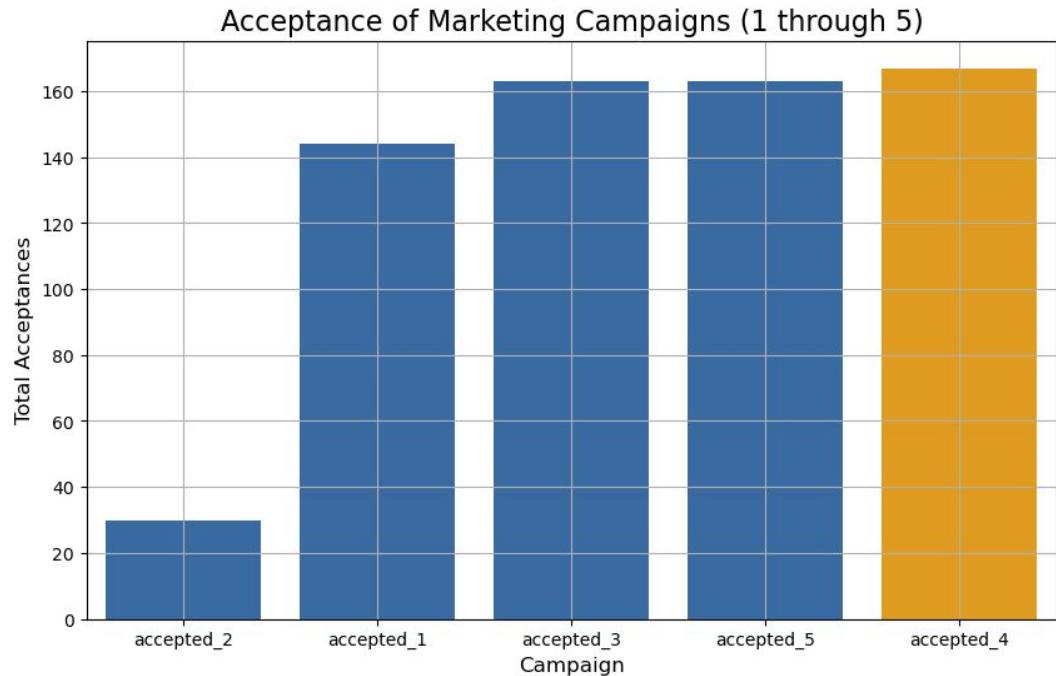


Marketing Campaign Performance

Chose a simple bar graph to easily compare the success of each marketing campaign.

Uses a SUM function to sum the total number of 1's or acceptances from each campaign.

Campaign number 4 was the most accepted offer at 167.



Catalog Purchases & Meat Regression

Each \$1 spent through catalog purchases is associated with **an additional \$55.89 spent on meat.**

The model **explains 52.4% of the variation** in meat spending patterns (slightly better than wine).

Baseline meat spending (without catalog influence) is **about \$18.16.**

OLS Regression Results						
Dep. Variable:	meat	R-squared:			0.524	
Model:	OLS	Adj. R-squared:			0.524	
Method:	Least Squares	F-statistic:			2463.	
Date:	Tue, 22 Oct 2024	Prob (F-statistic):			0.00	
Time:	13:43:33	Log-Likelihood:			-14486.	
No. Observations:	2240	AIC:			2.898e+04	
Df Residuals:	2238	BIC:			2.899e+04	
Df Model:	1					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
const	18.1617	4.452	4.079	0.000	9.431	26.892
catalog_purchases	55.8923	1.126	49.628	0.000	53.684	58.101
Omnibus:	915.647	Durbin-Watson:			1.935	
Prob(Omnibus):	0.000	Jarque-Bera (JB):			8291.508	
Skew:	1.680	Prob(JB):			0.00	
Kurtosis:	11.806	Cond. No.			5.51	

Catalog Purchases & Wine Regression

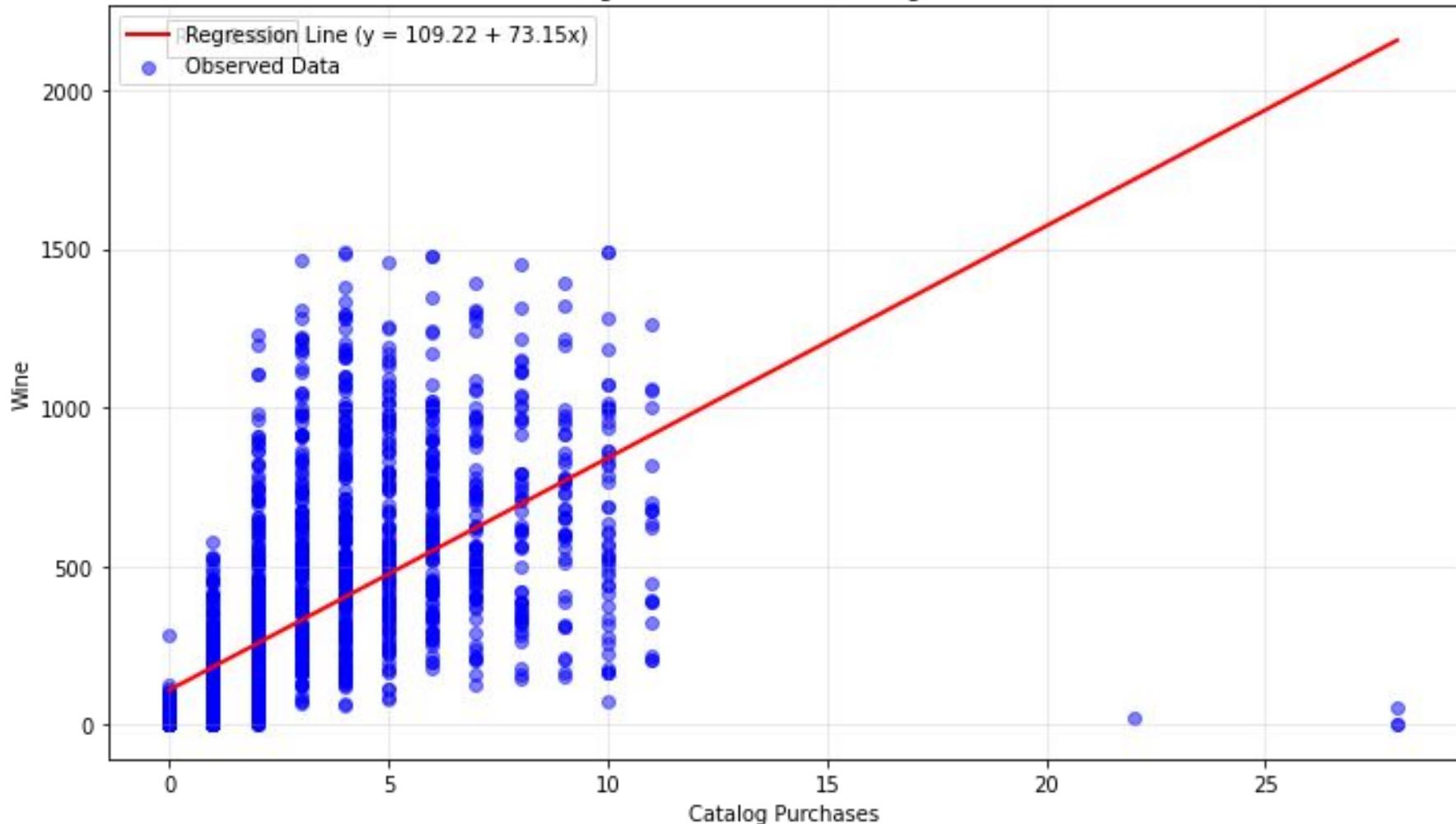
Each \$1 spent through catalog purchases is associated with an **additional \$73.15 spent on wine.**

The model **explains 40.4% of the variation in wine** spending patterns.

Baseline wine spending (without catalog influence) is **about \$109.22.**

OLS Regression Results						
Dep. Variable:	wine	R-squared:	0.404			
Model:	OLS	Adj. R-squared:	0.403			
Method:	Least Squares	F-statistic:	1514.			
Date:	Tue, 22 Oct 2024	Prob (F-statistic):	2.10e-253			
Time:	13:44:54	Log-Likelihood:	-15634.			
No. Observations:	2240	AIC:	3.127e+04			
Df Residuals:	2238	BIC:	3.128e+04			
Df Model:	1					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
const	109.2151	7.432	14.696	0.000	94.642	123.789
catalog_purchases	73.1468	1.880	38.910	0.000	69.460	76.833
Omnibus:	404.874	Durbin-Watson:	1.978			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	6922.655			
Skew:	0.325	Prob(JB):	0.00			
Kurtosis:	11.588	Cond. No.	5.51			

OLS Regression: Wine vs Catalog Purchases



Implications



Enhanced Decision Making

Making **data-driven decisions** to create more effective promotional strategies. For example, **promote wine and meat on the front page** of the catalog.



Increased ROI

Maximizing returns on marketing investments by focusing on **high potential customer segments** by advertising the right products.



Improved Customer Experience

Enhancing customer satisfaction and loyalty through **targeted product promotions** interesting using the **proper marketing channel**.





04 Conclusion

n

Implications & Future
Extension

Future Extension



Developing Predictive Models

Incorporate machine learning techniques like classification models that further predict customer responses to promotions.



Customer Segmentation Analysis

Apply clustering techniques to segment customers based on their purchasing behaviors.



A/B Testing Framework

Design A/B tests that compare the effectiveness of different promotional strategies.



Thanks!



Any question?

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, infographics & images by **Freepik**

Please keep this slide for attribution





05

Appendix

Contributions to Slides

Elaina: data collection, correlation matrix, histogram, box plot, and bar graph, future extension

Jacob: Problem statement, importance of the problem, regression analysis and business implications

Work Cited

<https://www.kaggle.com/datasets/rodsaldanha/arketing-campaign>