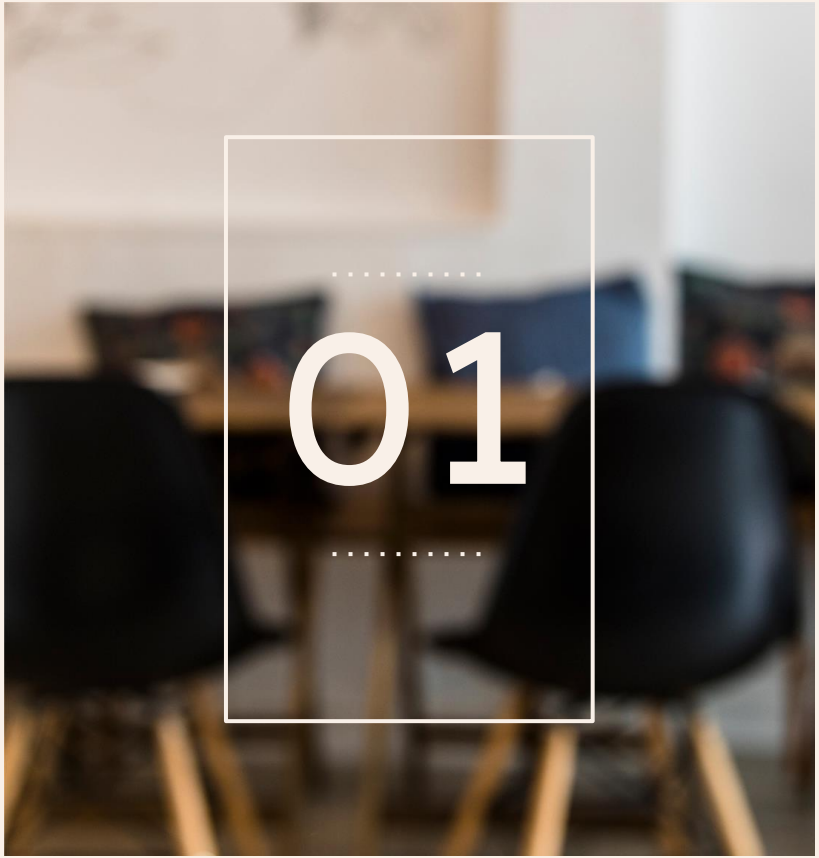
# Maven Rewards Challenge

Mohammadhadi Shahhosseini




# Overview

# Challenge Objective

In this Challenge, we'll play the role of a Sr. Marketing Analyst at Maven Cafe.

We've just run a test by sending different combinations of promotional offers to existing rewards members. Now that the 30-day period for the test has concluded, the task is to identify key customer segments and develop a data-driven strategy for future promotional messaging & targeting.


# ABOUT Data Set

Data that simulates the behavior of Cafe Rewards members over a 30-day period, including their transactions and responses to promotional offers. The data is contained in three files: one with details on each offer, another with demographic information on each customer, and a third with the activity for each customer during the period. The activities are divided into offer received, offer viewed, offer accepted, and transaction. For a transaction to be attributed to an offer, it must occur at the same time as when the offer was "completed" by the customer.

# Tables

| | Field | Description |
|---|---|---|
| Offers | offer_id | Unique offer ID (primary key) |
| | offer_type | type of offer: bogo (buy one, get one), discount, or informational |
| | difficulty | minimum amount required to spend in order to be able to complete the offer |
| | reward | reward (in dollars) obtained by completing the offer |
| | duration | days a customer has to complete the offer once they have received it |
| | channels | list of marketing channels used to send the offer to customers |

# Tables

| | Field | Description |
|---|---|---|
| Customers | customer_id | Unique customer ID (primary key) |
| | became_member_on | Date when the customer created their account (yyyymmdd) |
| | gender | Customer's gender: (M)ale, (F)emale, or (O)ther |
| | age | Customer's age |
| | income | Customer's estimated annual income, in USD |

# Tables

| | Field | Description |
|---|---|---|
| Events | customer_id | Customer the event is associated with (foreign key) |
| | event | Description of the event (transaction, offer received, offer viewed, or offer completed) |
| | value | Dictionary of values associated with the event (amount for transactions, offer_id for offers received and viewed, and offer_id & reward for offers completed). |
| | time | Hours passed in the 30-day period (starting at 0) |

# 02

# Data Preparation & Feature Engineering

Offers

Customers

Events

Merge

Final Dataset

## Data Cleaning

- Extracted relevant fields (e.g., offer_id, reward, amount) from JSON-like structures.
- Corrected data types.
- Removed rows with missing income.

## Feature Engineering

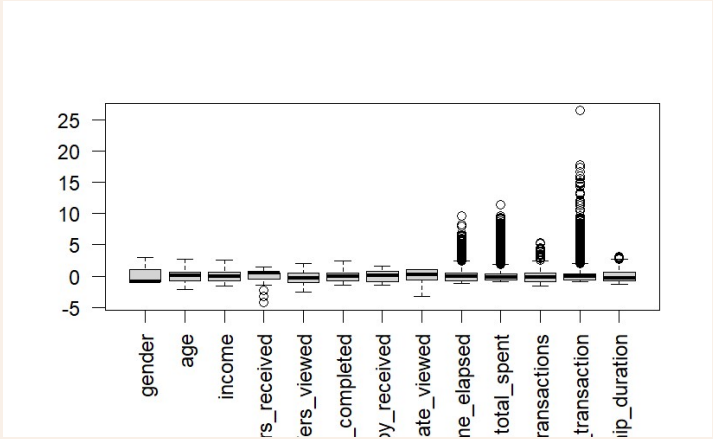
- Calculated features like membership_duration, time_elapsed, total_spent, num_transactions, and offer completion rates.
- Created lagged variables to capture time between offer receipt and completion.
- Merged customer interaction data (received/completed offers, transaction data).

# ANALYSIS

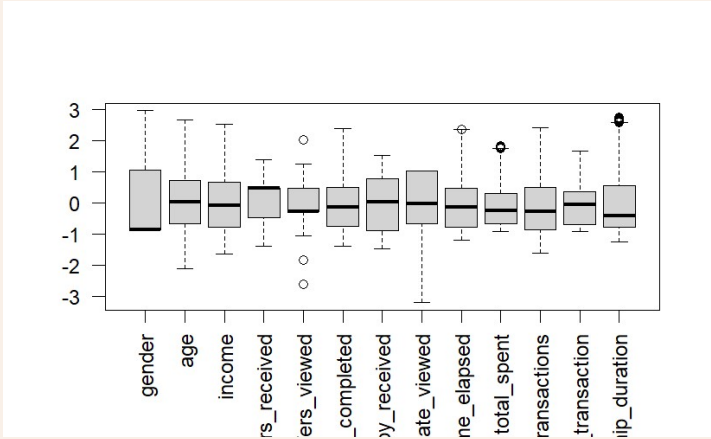
Unsupervised Customer Segmentation

# Data Preparation & Outlier Detection

- Encoded categorical data (e.g., gender) to numeric.
- Normalized data for analysis.
- Created membership_duration feature.
- Identified and removed outliers using IQR method.

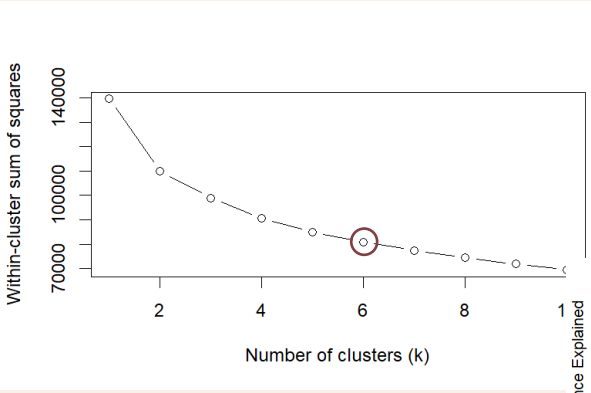


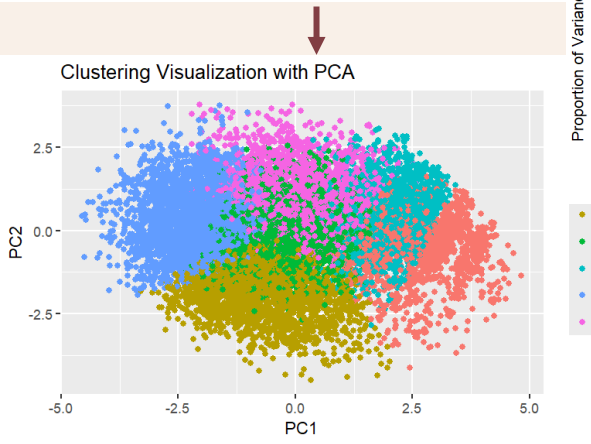

Removing about 2000 outliers



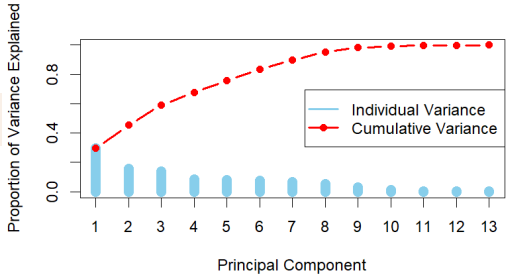
# K-Means clustering on original and PCA data.

## Original Data



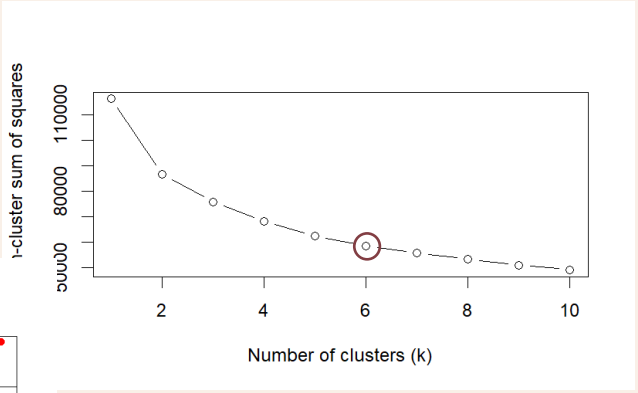



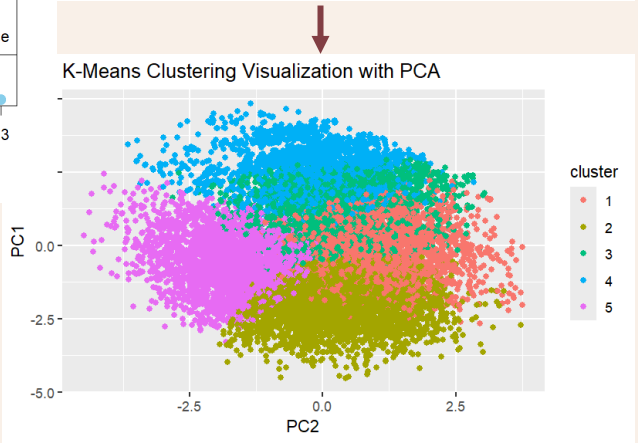

**PCA reduced dimensions: top 6 components explain 84% variance.**




**Elbow method used to determine k = 6 clusters.**

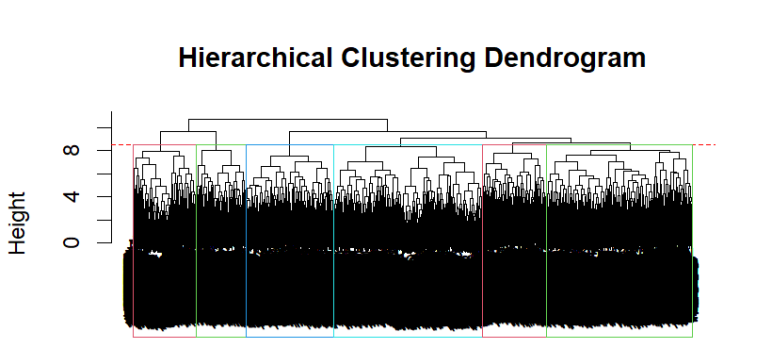
## PCA Data



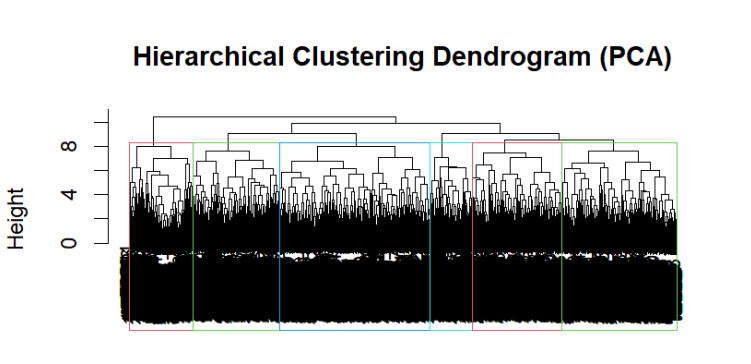



# K-Means clustering on original and PCA data.

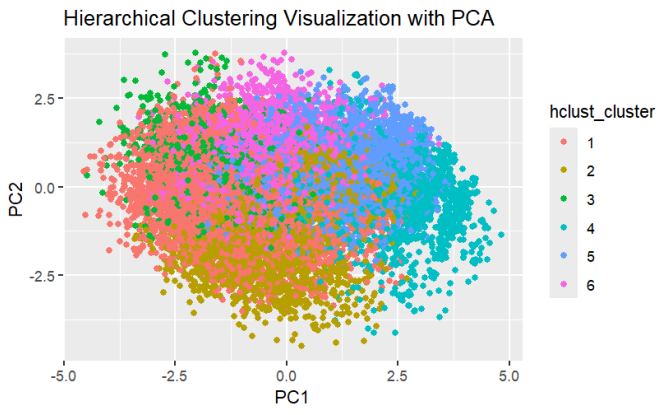
Original Data

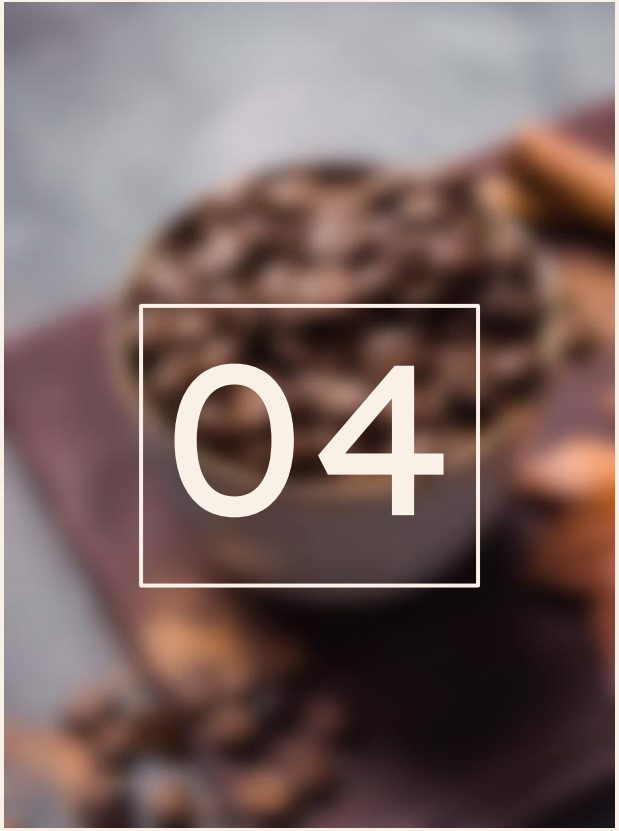



PCA Data




**Dendrogram visualization with 6 clusters.**






# ANALYSIS

Supervised Offer Personalization

# Overview Flowchart




### Data Loading & Preparation

- Load Data & Handle Missing Values
- Convert Data Types

### Feature Engineering

- Create Features (e.g., Membership Duration)
- Encode Categorical Variables (One-Hot /Ordinal)

### Modeling

- Train XGBoost Model
- Handle Class Imbalance (Oversampling /Weights)

### Evaluation

- Evaluate Model (Precision, Recall, F1, Accuracy)

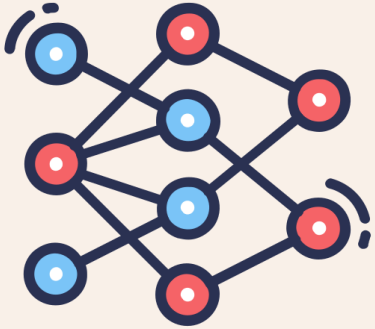
# Model Building - XGBoost

## Data Split

70% Train

30% Test

## Parameters

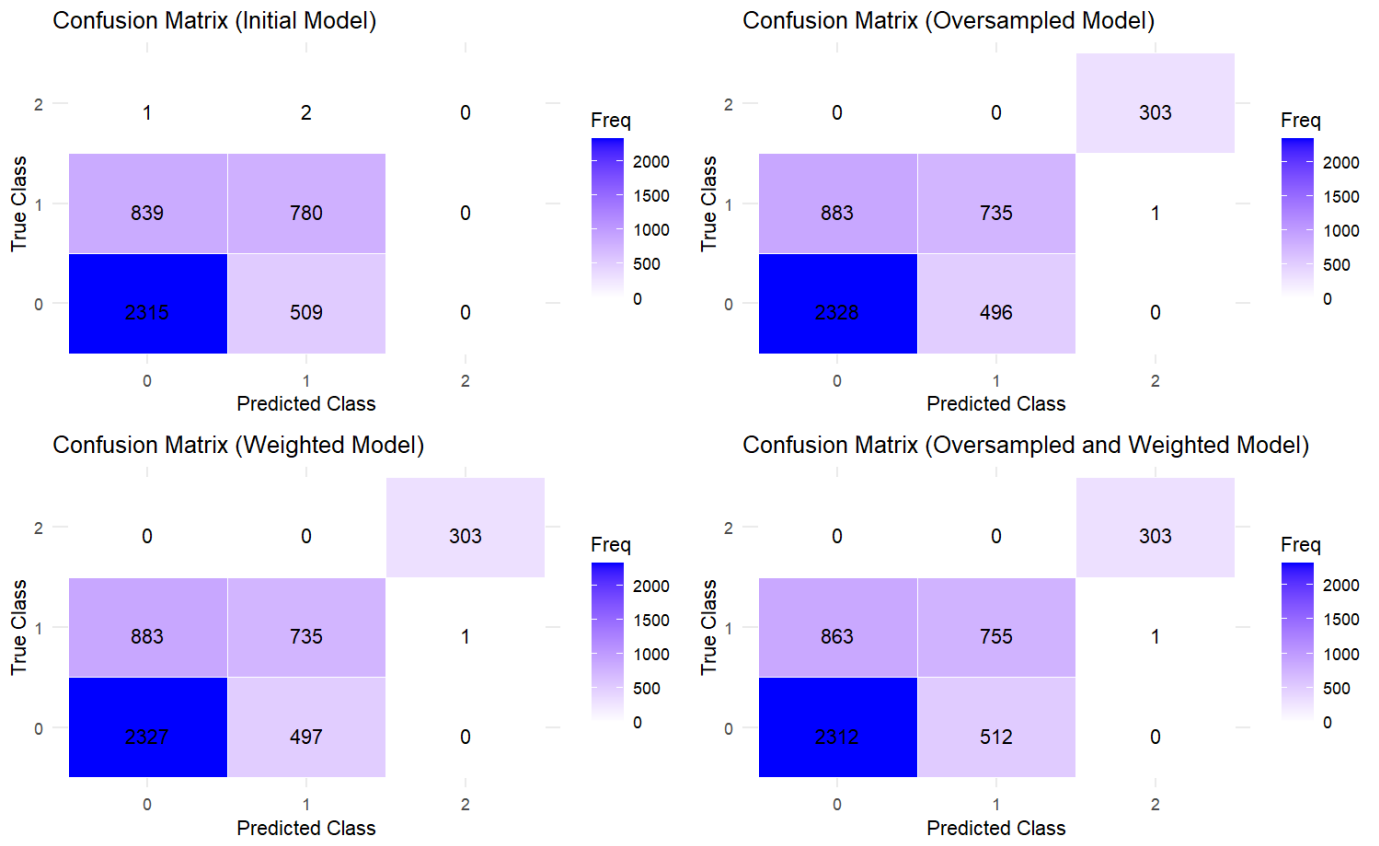


- Learning rate (eta) = 0.3
- Max depth = 6
- Evaluation metric: mlogloss

## Handling Bias

- Applied **class weights** to account for imbalance.
- Implemented **oversampling** for the minority class (informational offers).

# Results and Metrics



# Results and Metrics

| | | Class 1 | Class 2 | Class 3 |
|---|---|---|---|---|
| Initial Model **69 %** | **Precision** | 0.7338 | 0.6042 | NA |
| | **Recall** | 0.8198 | 0.4818 | 0.0000000 |
| | **F1 - Score** | 0.7744 | 0.5361 | NA |
| Weighted Model **69 %** | **Precision** | 0.7316 | 0.6067 | NA |
| | **Recall** | 0.8251 | 0.4725 | 0.0000000 |
| | **F1 - Score** | 0.7755 | 0.5312 | NA |
| Oversampled Model **70.9 %** | **Precision** | 0.7249 | 0.5966 | 0.99671 |
| | **Recall** | 0.8240 | 0.4540 | 1.00000 |
| | **F1 - Score** | 0.7713 | 0.5156 | 0.99835 |
| Oversampled and Weighted Model **71 %** | **Precision** | 0.728189 | 0.5958958 | 0.9967105 |
| | **Recall** | 0.8186969 | 0.4663372 | 1 |
| | **F1 - Score** | 0.7707951 | 0.5232155 | 0.9983526 |

# THANKS