



STARBUCKS CUSTOMER REWARD ANALYSIS

TEAM SIXES



ABOUT US



Bac Le



Evelyn Dong



Xue Chen



Ye Wang



Zhuomin Ni

EXECUTIVE SUMMARY



STUDY STARBUCKS MOBILE APP CUSTOMER DATA

Study a dataset from 14.8K Starbucks users to understand their purchase behavior when it comes to promotions



K-MEANS CLUSTERING ANALYSIS

Segment Starbucks customers into 4 clusters and study the performance of each offer on these clusters



RECOMMENDATIONS

Identifying influential factors, tailoring offers to each cluster, converting low-value cluster to high value one

AGENDA

1

Problem Statement

2

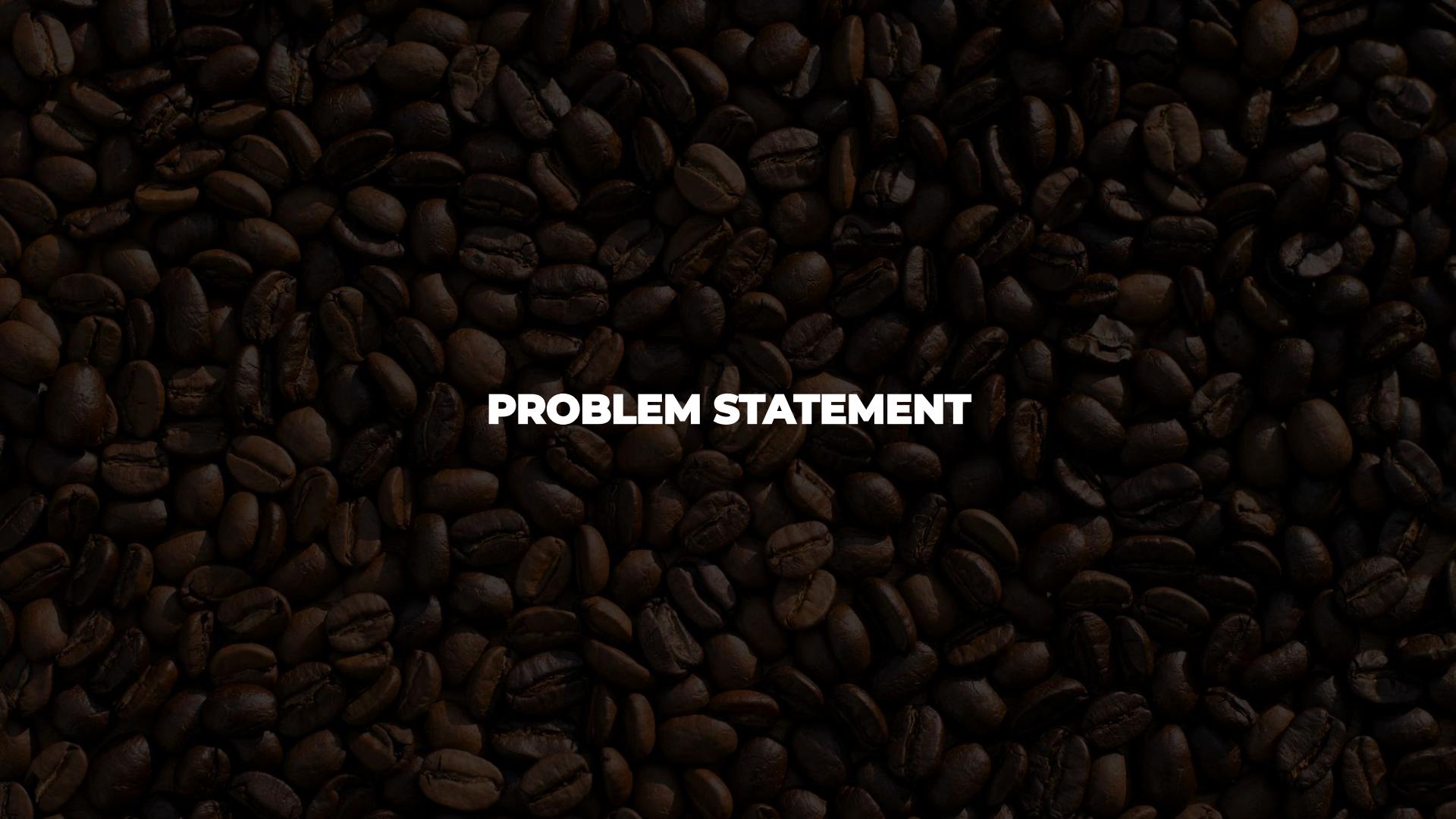
Data & Methodology

3

Results & Insights

4

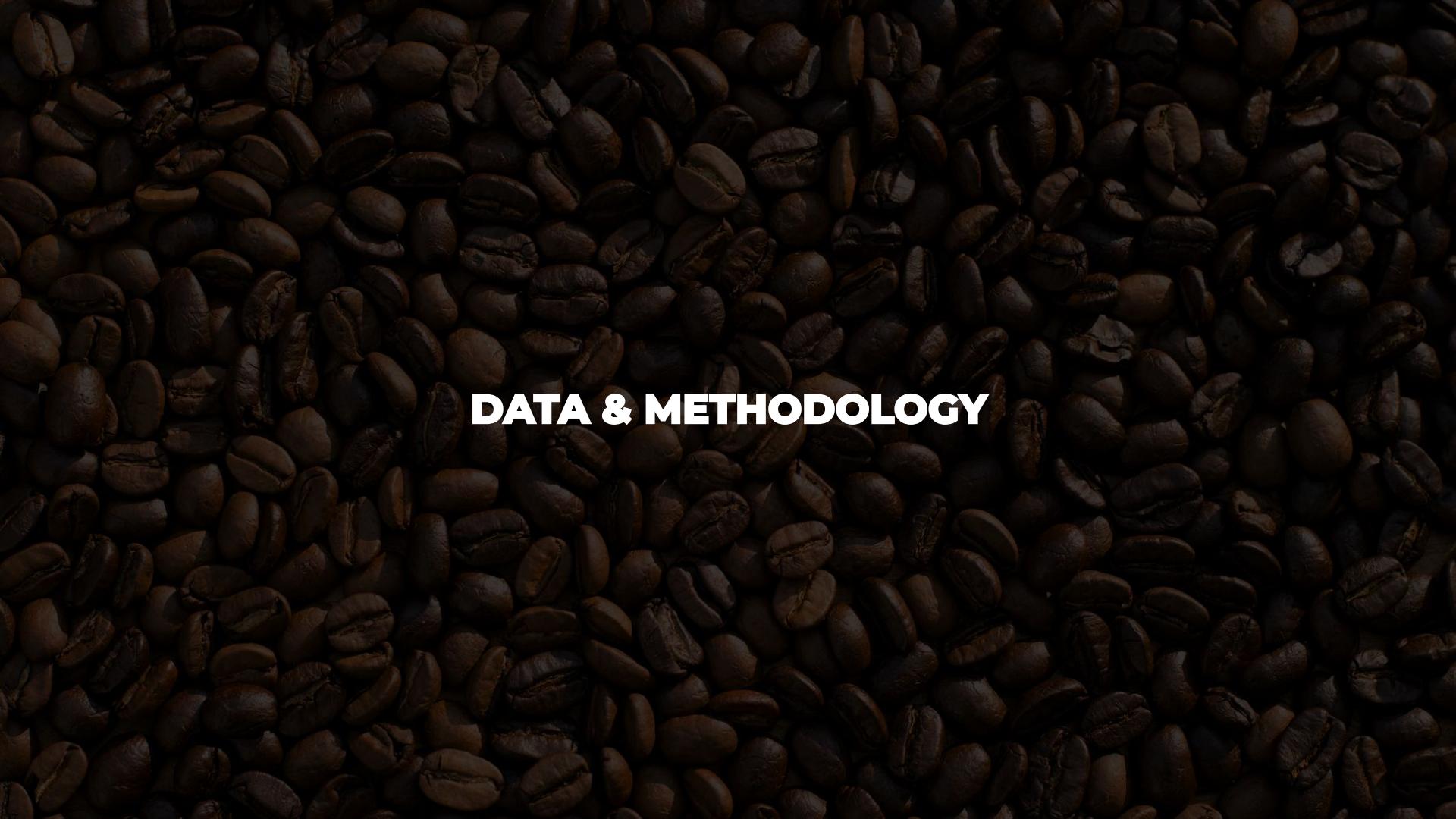
Recommendations & Managerial Implications



PROBLEM STATEMENT



**Identifying high-performing customer segments
and promotion offers on Starbucks Mobile App**



DATA & METHODOLOGY

DATA & METHODOLOGY

DATASET DESCRIPTION

DATASET Starbucks app customer rewards program

SOURCE <https://www.kaggle.com/blacktile/starbucks-app-customer-reward-program-data>

CONTEXT The dataset was provided to Udacity scholars by Starbucks for research projects

CONTENT Including 3 files:

portfolio.json (1.56KB) - containing offer ids and offer metadata (type, reward, duration, etc.)

profile.json (1.92MB) - containing demographic data for each customer

transcript.json (38.38MB) - containing records for transactions, offers received, offers viewed, and offers completed

DATA & METHODOLOGY

OVERVIEW OF USER DATA

- `became_member_on` (int) — date when customer created an app account
- `gender` (str) — gender of the customer
- `id` (str) — customer id
- `income` (float) — customer's income

	<code>gender</code>	<code>age</code>	<code>id</code>	<code>became_member_on</code>	<code>income</code>
0	None	118	68be06ca386d4c31939f3a4f0e3dd783	20170212	NaN
1	F	55	0610b486422d4921ae7d2bf64640c50b	20170715	112000.0
2	None	118	38fe809add3b4fcf9315a9694bb96ff5	20180712	NaN
3	F	75	78afa995795e4d85b5d9ceeca43f5fef	20170509	100000.0
4	None	118	a03223e636434f42ac4c3df47e8bac43	20170804	NaN
...

DATA & METHODOLOGY

OVERVIEW OF TRANSACTION DATA

- event (str) — record description (ie transaction, offer received, offer viewed, etc.)
- person (str) — customer id
- time (int) — time in hours since start of test. The data begins at time t=0
- value — (dict of strings) — either an offer id or transaction amount depending on the record

person	event	value	time
78afa995795e4d85b5d9ceeca43f5fef	offer received	{'offer id': '9b98b8c7a33c4b65b9aebfe6a799e6d9'}	0
78afa995795e4d85b5d9ceeca43f5fef	offer viewed	{'offer id': '9b98b8c7a33c4b65b9aebfe6a799e6d9'}	6
78afa995795e4d85b5d9ceeca43f5fef	transaction	{'amount': 19.89}	132
78afa995795e4d85b5d9ceeca43f5fef	offer completed	{'offer_id': '9b98b8c7a33c4b65b9aebfe6a799e6d9...'	132
78afa995795e4d85b5d9ceeca43f5fef	transaction	{'amount': 17.78}	144

DATA & METHODOLOGY

METHODOLOGY: K-MEANS CLUSTERING

I Input Variables

A. Demographic variables

1. Exclude the age variable
2. Calculate membership duration for each customer
3. Drop NaN values
4. 1 = Female, 0 = Male

B. Behavioral variables

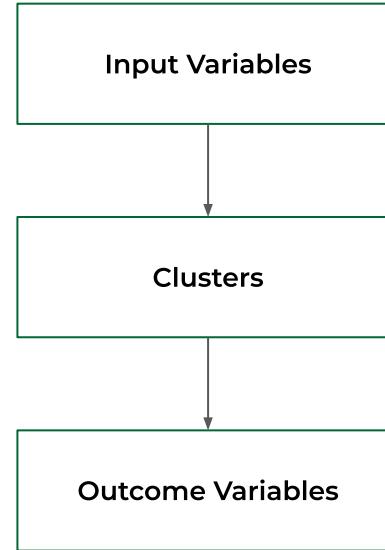
1. Count the frequency of user activity
 - # of offer completed
 - # of offer received
 - # of offer viewed
 - # of transaction
2. Sum the total transaction amount

II

Clusters: 4 clusters representing 4 customer personas

III

Outcome Variables: Completion rate for each offer type



DATA & METHODOLOGY

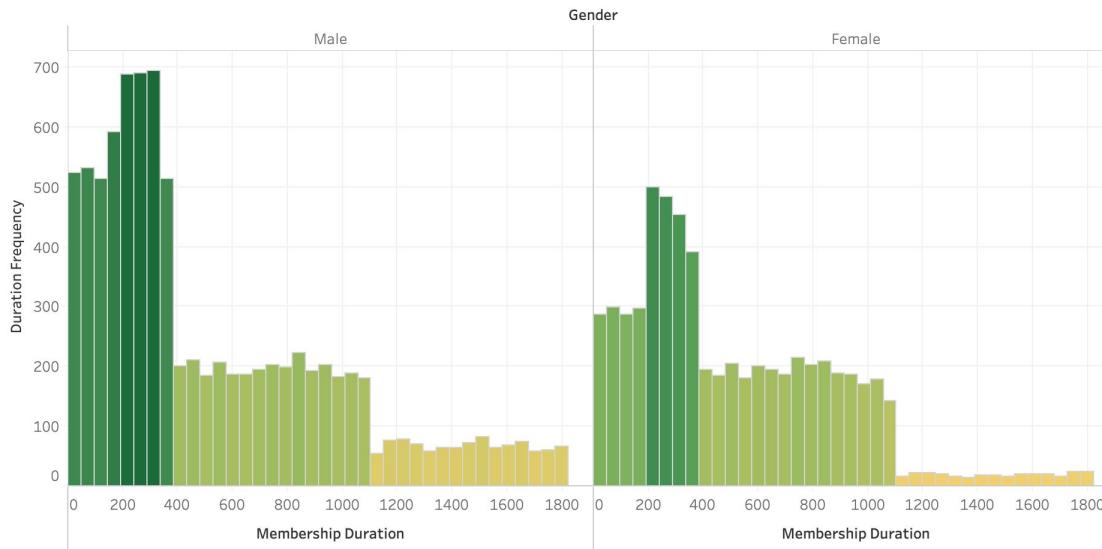
INPUT VARIABLES

14825 users (8696 male, 6129 female)

DATA & METHODOLOGY

INPUT VARIABLES

14825 users (8696 male, 6129 female)

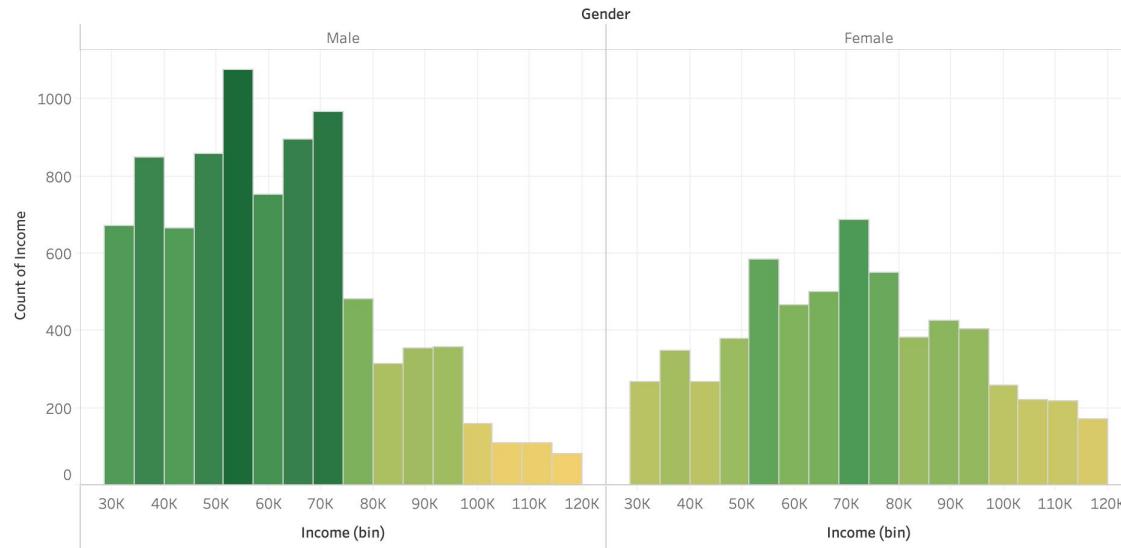


Duration is right-skewed for both gender groups

DATA & METHODOLOGY

INPUT VARIABLES

14825 users (8696 male, 6129 female)



Income is left skewed for male, but normal distributed for female

DATA & METHODOLOGY

DRAWBACKS OF THE METHODOLOGY

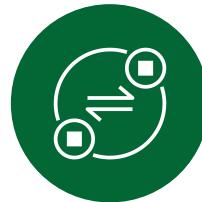


Transactions ≠ redeemed offers

No 'activate the offer' option
in the app

Transaction happens without
users having viewed an offer

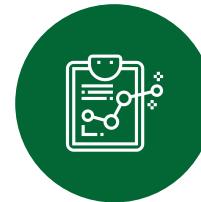
We don't know which
transaction is related to the
completion of which offer



Correlation between offers

Users receive different
promotions at varied
frequencies across time

We don't know the correlation
between offers (i.e. whether a
BOGO offer influences a
discount offer)



Assumptions about the data

Data privacy of age

Naive assumptions about
membership duration

Each user doesn't have the
equal chance to receive an
offer



RESULTS & INSIGHTS

RESULTS & INSIGHTS

CLUSTER RESULTS

cluster	gender	income	duration	offer received	offer viewed	offer completed	transaction	value	count
1	0.1847982709	51013.32853	1025.573127	4.362391931	3.167867435	2.109149856	14.00648415	94.10608429	2776
2	0.17728471	54202.32865	204.1509227	4.515817223	3.16586116	0.816344464	4.982205624	31.54453866	4552
3	0.5407515861	73086.62762	570.6293314	5.210102489	4.415812592	3.927769644	10.55270864	228.1650415	4098
4	0.7628714328	82900.26478	479.8546631	3.672844954	2.518093557	1.993821712	5.634304207	116.2379612	3399

Cluster 1 - Loyal male customers who live on coffee

Cluster 2 - Occasional male customers

Cluster 3 - Loyal high-value customers

Cluster 4 - Occasional high-value female customers

RESULTS & INSIGHTS

CLUSTER RESULTS

cluster	gender	income	duration	offer received	offer viewed	offer completed	transaction	value	count
1	0.1847982709	51013.32853	1025.573127	4.362391931	3.167867435	2.109149856	14.00648415	94.10608429	2776
2	0.17728471	54202.32865	204.1509227	4.515817223	3.16586116	0.816344464	4.982205624	31.54453866	4552
3	0.5407515861	73086.62762	570.6293314	5.210102489	4.415812592	3.927769644	10.55270864	228.1650415	4098
4	0.7628714328	82900.26478	479.8546631	3.672844954	2.518093557	1.993821712	5.634304207	116.2379612	3399

Cluster 1 - Loyal male customers who live on coffee

Male-dominant

Medium income

Long-time customers

High offer view rate (73%), average offer completion rate (48%)

High volume of transactions, low transaction value (\$6.71 per transaction)

RESULTS & INSIGHTS

CLUSTER RESULTS

cluster	gender	income	duration	offer received	offer viewed	offer completed	transaction	value	count
1	0.1847982709	51013.32853	1025.573127	4.362391931	3.167867435	2.109149856	14.00648415	94.10608429	2776
2	0.17728471	54202.32865	204.1509227	4.515817223	3.16586116	0.816344464	4.982205624	31.54453866	4552
3	0.5407515861	73086.62762	570.6293314	5.210102489	4.415812592	3.927769644	10.55270864	228.1650415	4098
4	0.7628714328	82900.26478	479.8546631	3.672844954	2.518093557	1.993821712	5.634304207	116.2379612	3399

Cluster 2 - Occasional male customers

Male-dominant

Medium income

New customers

High offer view rate (70%), low offer completion rate (18%)

Low volume of transactions, low transaction value (\$6.33 per transaction)

RESULTS & INSIGHTS

CLUSTER RESULTS

cluster	gender	income	duration	offer received	offer viewed	offer completed	transaction	value	count
1	0.1847982709	51013.32853	1025.573127	4.362391931	3.167867435	2.109149856	14.00648415	94.10608429	2776
2	0.17728471	54202.32865	204.1509227	4.515817223	3.16586116	0.816344464	4.982205624	31.54453866	4552
3	0.5407515861	73086.62762	570.6293314	5.210102489	4.415812592	3.927769644	10.55270864	228.1650415	4098
4	0.7628714328	82900.26478	479.8546631	3.672844954	2.518093557	1.993821712	5.634304207	116.2379612	3399

Cluster 3 - Loyal high-value customers

Neutral gender

High income

Long-time customers

High offer view rate (85%), high offer completion rate (75%)

High volume of transactions, high transaction value (\$21.62 per transaction)

RESULTS & INSIGHTS

CLUSTER RESULTS

cluster	gender	income	duration	offer received	offer viewed	offer completed	transaction	value	count
1	0.1847982709	51013.32853	1025.573127	4.362391931	3.167867435	2.109149856	14.00648415	94.10608429	2776
2	0.17728471	54202.32865	204.1509227	4.515817223	3.16586116	0.816344464	4.982205624	31.54453866	4552
3	0.5407515861	73086.62762	570.6293314	5.210102489	4.415812592	3.927769644	10.55270864	228.1650415	4098
4	0.7628714328	82900.26478	479.8546631	3.672844954	2.518093557	1.993821712	5.634304207	116.2379612	3399

Cluster 4 - Occasional high-value female customers

Female-dominant

High income

Long-time customers

High offer view rate (69%), average offer completion rate (54%)

Low volume of transactions, high transaction value (\$20.63 per transaction)

RESULTS & INSIGHTS

OFFERS' DIFFERENTIATING FACTORS



Offer Types

BOGO
Discount
Informational



Minimum Required Spending

Minimum amount to complete an offer



Reward

Amount received for completing an offer



Duration

How long an offer is available for redeeming



Distribution Channels

Mobile
Email
Web
Social

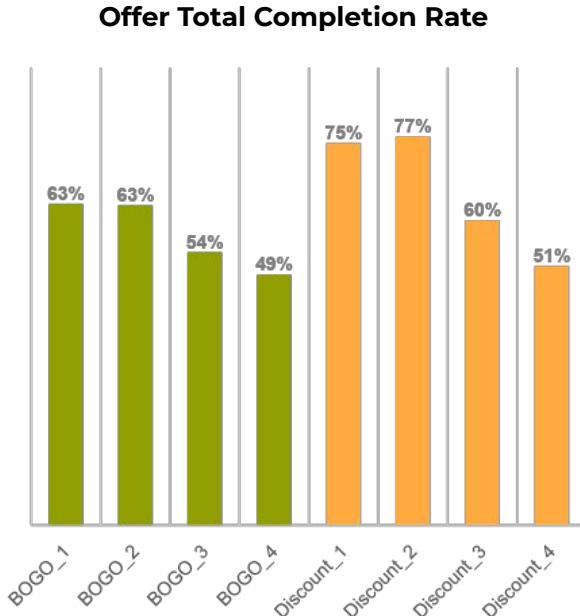
RESULTS & INSIGHTS

OFFERS SUMMARY

Offer Type	Minimum Spend (\$)	Reward Earned (\$)	Duration (Days)	Distribution Channels	Offer Distributed
BOGO_1	5	5	7	Web, Email, Mobile	7677
BOGO_2	5	5	5	Web, Email, Mobile, Social	7571
BOGO_3	10	10	7	Email, Mobile, Social	7658
BOGO_4	10	10	5	Web, Email, Mobile, Social	7593
Discount_1	7	3	7	Web, Email, Mobile, Social	7646
Discount_2	10	2	10	Web, Email, Mobile, Social	7597
Discount_3	10	2	7	Web, Email, Mobile	7632
Discount_4	20	5	10	Web, Email	7668
Informational_1	0	0	4	Web, Email, Mobile	7617
Informational_2	0	0	3	Email, Mobile, Social	7618

RESULTS & INSIGHTS

OFFERS PERFORMANCE

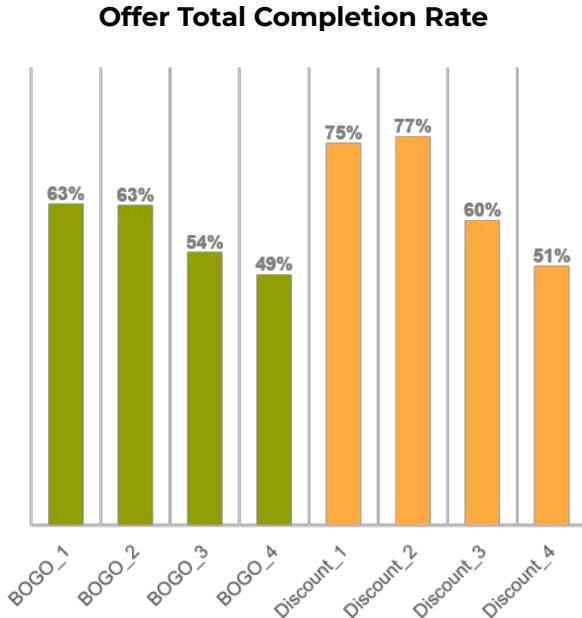


Offer Type	Minimum Spend (\$)	Reward Earned (\$)	Duration (Days)	Distribution Channels
BOGO_1	5	5	7	Web, Email, Mobile
BOGO_2	5	5	5	Web, Email, Mobile, Social
BOGO_3	10	10	7	Email, Mobile, Social
BOGO_4	10	10	5	Web, Email, Mobile, Social
Discount_1	7	3	7	Web, Email, Mobile, Social
Discount_2	10	2	10	Web, Email, Mobile, Social
Discount_3	10	2	7	Web, Email, Mobile
Discount_4	20	5	10	Web, Email

On average, Discount offers have higher completion rates than BOGO offers with similar promotion attributes.

RESULTS & INSIGHTS

OFFERS PERFORMANCE



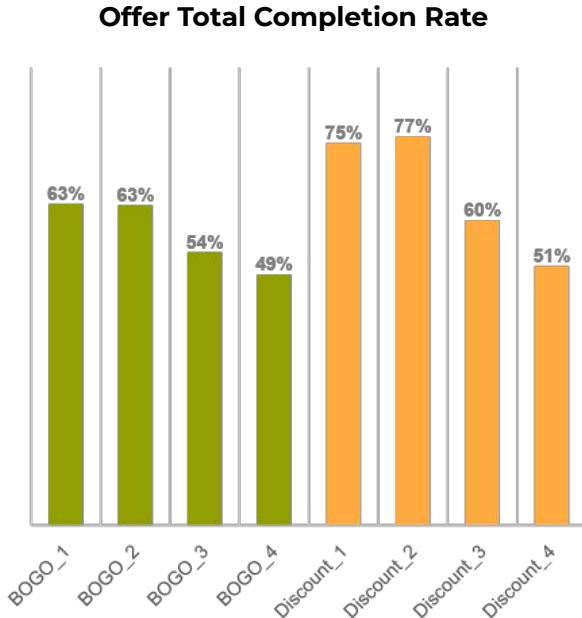
Offer Type	Minimum Spend (\$)	Reward Earned (\$)	Duration (Days)	Distribution Channels
BOGO_1	5	5	7	Web, Email, Mobile
BOGO_2	5	5	5	Web, Email, Mobile, Social
BOGO_3	10	10	7	Email, Mobile, Social
BOGO_4	10	10	5	Web, Email, Mobile, Social

BOGO_1&2 vs. BOGO_3&4: A LOWER minimum required spend encourages customers to redeem an offer.

BOGO_3 vs. BOGO_4: Both Distribution Channels and Duration may impact performance of offers.

RESULTS & INSIGHTS

OFFERS PERFORMANCE



Offer Type	Minimum Spend (\$)	Reward Earned (\$)	Duration (Days)	Distribution Channels
Discount_1	7	3	7	Web, Email, Mobile, Social
Discount_2	10	2	10	Web, Email, Mobile, Social
Discount_3	10	2	7	Web, Email, Mobile
Discount_4	20	5	10	Web, Email

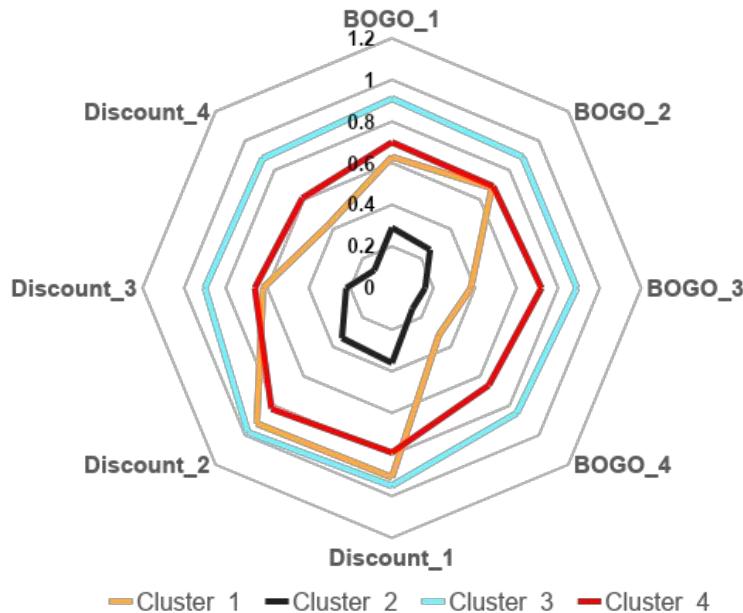
Discount_1~4: A LOWER minimum required spend encourages customers to redeem an offer.

Discount_2 vs. Discount_3: Distribution Channels and/or Duration may impact performance of offers.

RESULTS & INSIGHTS

OFFER POPULARITY ACROSS CUSTOMER SEGMENTS

Offer Completion Rate of Each Cluster



Cluster 1 - Loyal male customers who live on coffee

Strong preference for Discount offers

Cluster 2 - Occasional male customers

Strong preference for Discount offers but lower completion

Cluster 3 - Loyal high-value customers

Take advantage of every offer

Cluster 4 - Occasional high-value female customers

Attracted Least By Offers



RECOMMENDATIONS & MANAGERIAL IMPLICATIONS

RESULTS & INSIGHTS

RECOMMENDATION 1 - IDENTIFYING INFLUENTIAL FACTORS

Insights

When compare the offers, each two offers have at least two different attributes, not possible to determine which attribute relates to higher completion rate

Recommendation

A/B testing, one attribute different, others exactly the same

- BOGO vs. Discount,
- Short duration vs. Long duration
- Less Marketing channels vs. Full marketing channels

Managerial Implications

- What attributes can improve completion rate => Understand consumer preference => Develop better offers
- False negative/positive risk => Require large sample size

RESULTS & INSIGHTS

RECOMMENDATION 2 - TAILORING OFFERS TO CLUSTERS

Insights

Each cluster shows different pattern of preference on each type of offer.

Recommendation

Use STP model to give the clusters what they prefer

- Cluster 1: Discount 1+2, BOGO 1+2
- Cluster 3: Keep the current strategy
- Cluster 4: Less offer but more informational message for high income group

Managerial Implications

- Customized offer solutions to be more efficient on marketing promotion: higher overall completion rate
- Risk of cost increase on segmented promotion => ROI analysis

RESULTS & INSIGHTS

RECOMMENDATION 3 - CONVERTING CLUSTER 2 INTO CLUSTER 1

Insights

Cluster 1 and Cluster 2 share similar characteristics and preferences, and are only different in membership duration & purchasing behaviors. Therefore, there is a potential to convert Cluster 2 members into Cluster 1, which generates higher value for Starbucks.

Recommendation

- Research Cluster 2's preferences
- Develop special offers tailored to Cluster 2's preferences
- Convert Cluster 2 members (occasional) into Cluster 1 members (loyal)

Managerial Implications

- Increase Cluster 2's transaction volume and completion rate
- Increase customer loyalty of the Cluster 2
- Follow up to see if we should change the strategy after the conversion

RESULTS & INSIGHTS

SUCCESS MEASUREMENT



Cluster
Engagement



Completion Rate



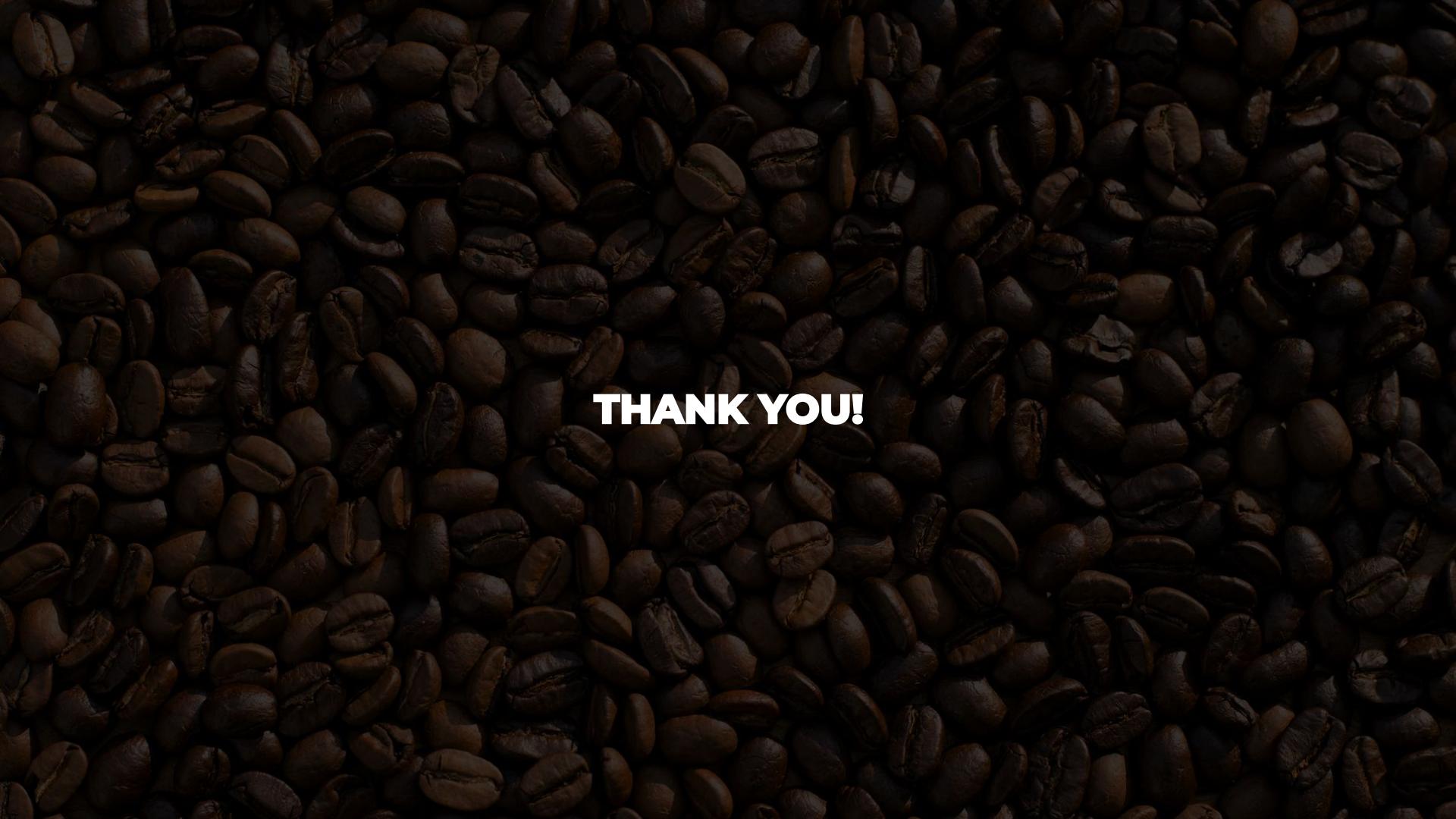
Return On
Investment



Sales



Customer
Retention

The background of the image is a dense, dark pile of coffee beans, creating a rich, textured pattern.

THANK YOU!