


Starbucks Rewards:

Analyzing Customers' Response to Promo Offers in Differentiating Loyalty Maximization

Non-Technical Presentation
Capstone 2 Project

by Yerusalem Gizaw

June 2021



Free Coffee is a Tap Away!

Let’s educate the Rewards Members

and

create a win-win situation for all!

How does Starbucks differentiate its loyalty maximization by measuring the members' response in the Rewards program from 07/2013 - 07/2018 to the personalized promo offers by January 2022?

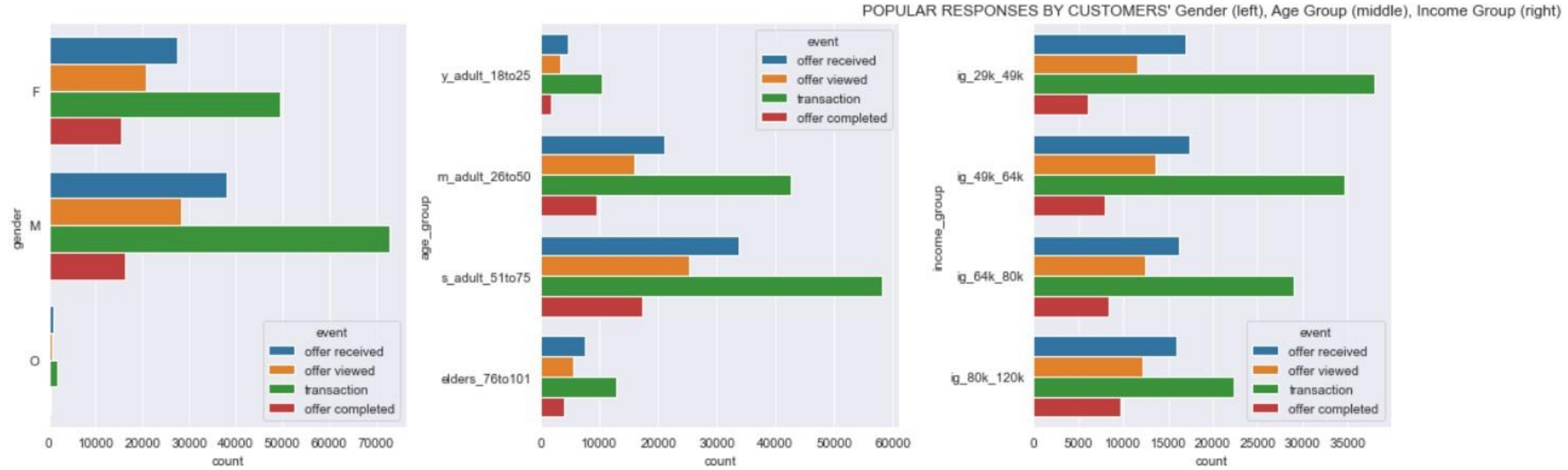
1) Do the Starbucks customers know why the Rewards program was created, and how does Starbucks measure the Reward members' response to offers? ➡	▪ Food & Drinks delivered with flexibility, convenience, and Rewards to thank loyal customers ▪ 44.8% of the personalized offers were only received but not completed by members
2) Who are the Rewards program participants, and which personalized promo offers did they complete the most? ➡	▪ 57.2% Males and 41.8% Females between average ages of 51 and 75 with income ranging b/n 49k-64K ▪ BOGO-43%, followed by Discount-42%, and Information-15%
3) How can the Starbucks improve the factors affecting the Reward members' responses? ➡	▪ 45% members received specialized rewards from 07/13-07/18 without paying money, but ▪ Focus on Channels, Offer Types, Duration and Reward Values needed to ⬆ offer completion

Data: *July 2013 – July 2018 CSV Files (portfolio.csv, profile.csv, transcript.csv)*
Source: *Kaggle https://www.kaggle.com/ihormuliar/starbucks-customer-data/metadata*

(1)

Do the Starbucks customers know why the Rewards program was created, and how does Starbucks measure the members' response to offers?

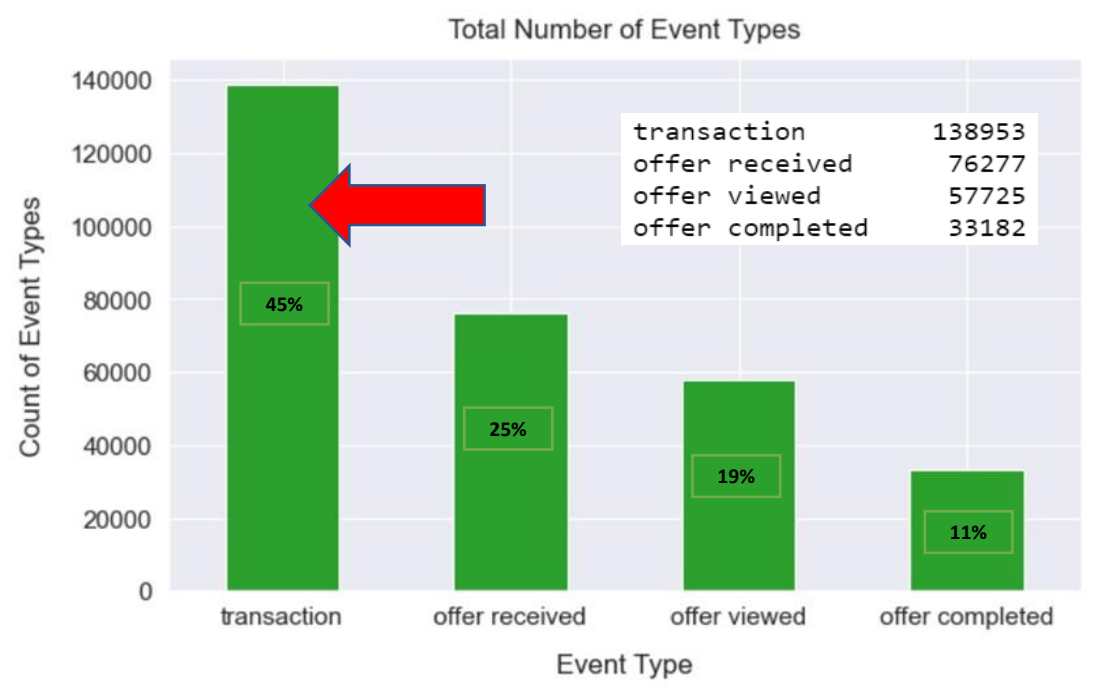
How do members respond to the personalized promo offers? What're the responses (events) to the offers and how are they related to one another?

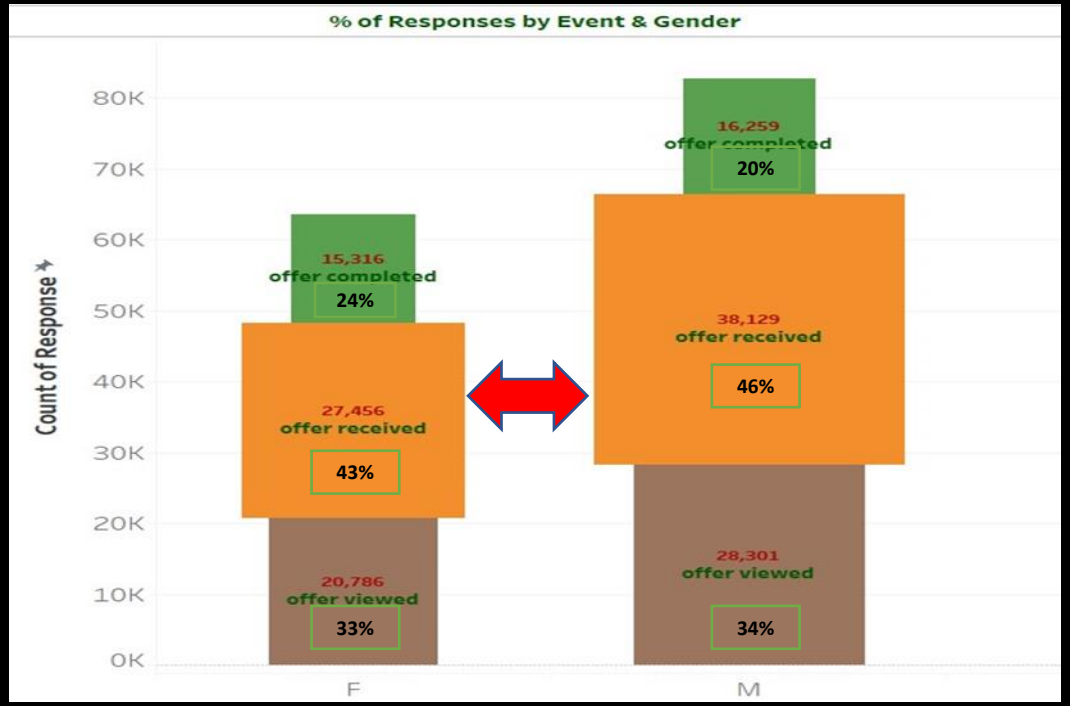


Key Insights

- Starbucks' Rewards program endeavors to thank loyal customers for patronizing its business and purchasing Starbucks products.
- Members respond to offers by receiving, viewing and completing offers to receive rewards.
- Transaction is the event with the highest numbers by all gender including the top users age-group of senior adults aged 51-75 yrs. as well as by adults 26-50 yrs.
- Transaction occurs when members make purchases by paying some money or $0 amount as a result of using their stars that are not directly/indirectly related to the promo offers, they received.

What does the total percentage of responses by offer type indicate across the male and female Reward members where responses to the transaction event is the highest?





Key Insights

- 45% members made transactions and received specialized rewards from 07/13-07/18 without paying money but no evidence if those transactions were attached to the personalized offers they received.
- 22% of promo offers get completed by all members.
- 44.8% of the personalized offers were only received but not viewed or completed by all members
- 33.5% of the personalized offers were only received and viewed but not completed by all members

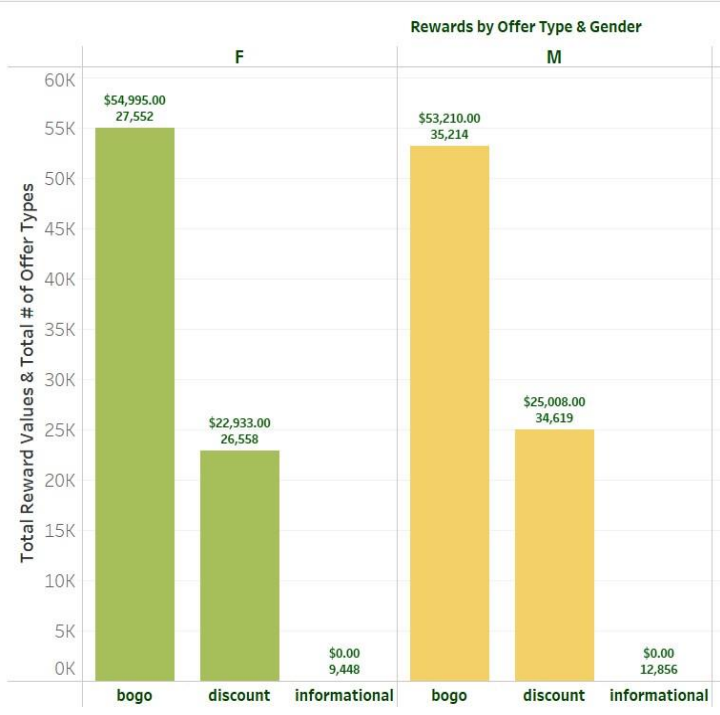
(2)
Who are the Rewards program participants, and which personalized promo offers did they complete the most?

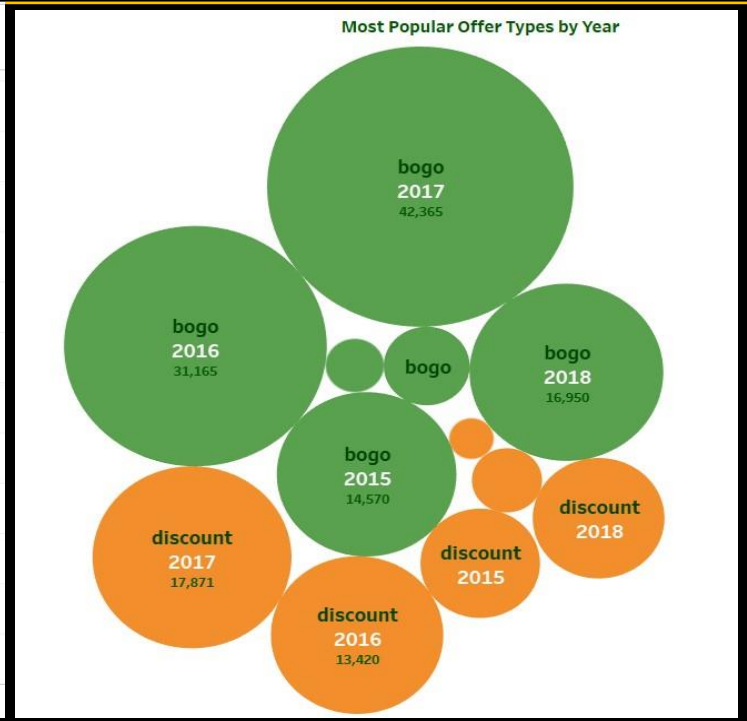
What is the demographic segmentation of over 17,000 customers who joined the Starbucks Rewards program during the 5 years at the participating stores? Which segment stand out at the top?

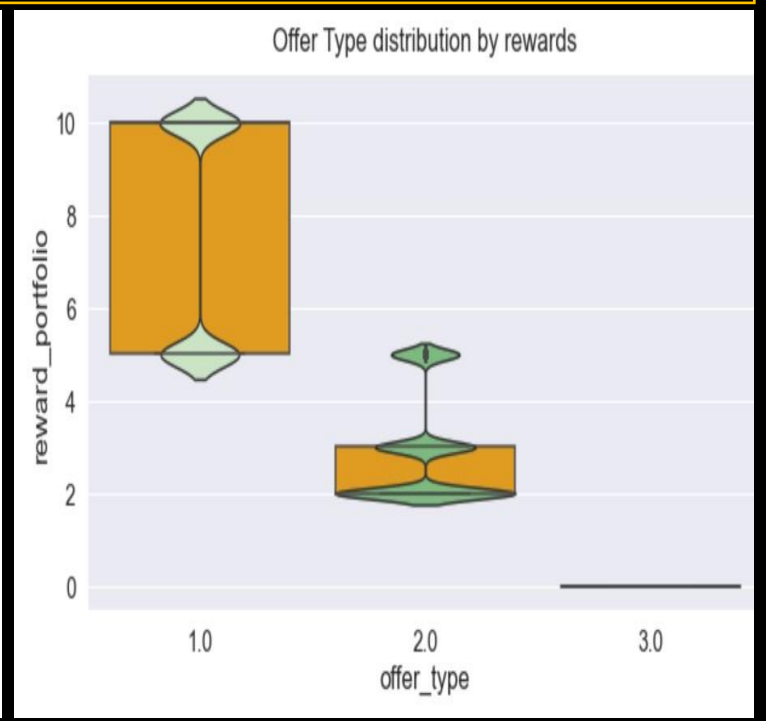
Key Insights

- **Gender**: Of the 17,000 Reward members enrolled = 57% male and 41% females (07/2017 – 07/2018)
- **Age:** 51-75 age group make up the 50% (7482), followed by the 26-50 yrs = 32% (4681), 74-101 yrs = 11% (1650), 18-24 yrs=7%
- **Income:** The difference in income distribution didn't show statistically significant impact where income group 49k-64k and 29k-49k make up 26% members and income group 64k-80k and 80k-120k = 24%

While both BOGO and DISCOUNT are almost equally popular, why do we see a difference between their rewards $ value and how was that observed over the 5 years period? Which year stand out the most?





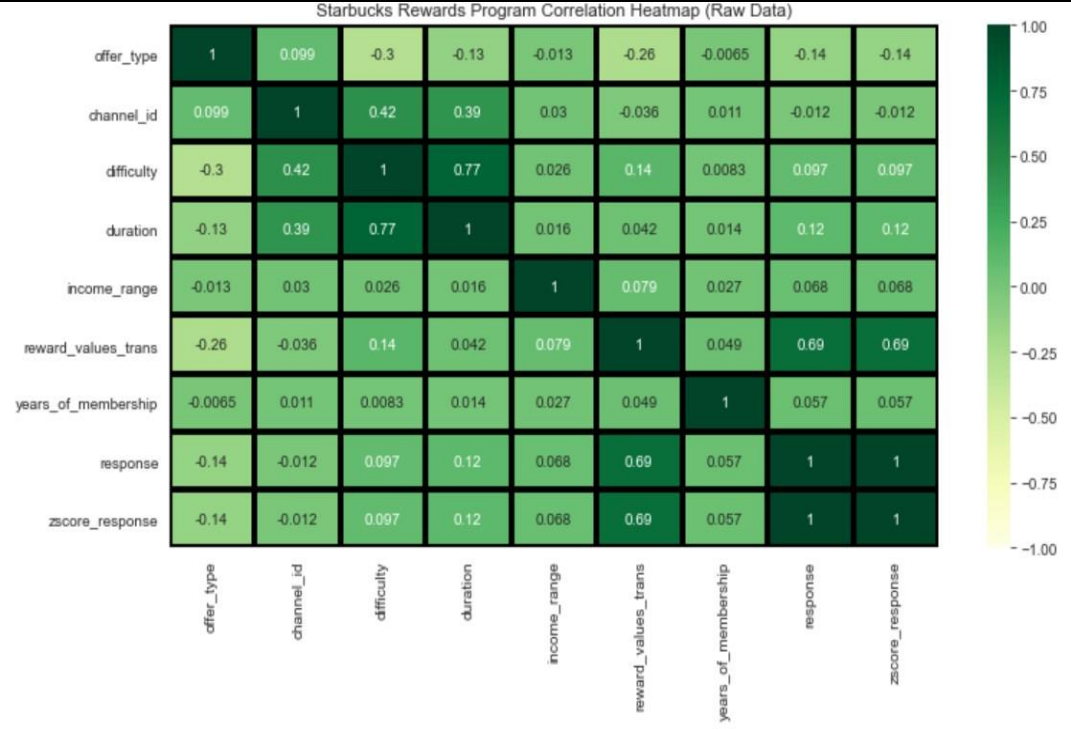


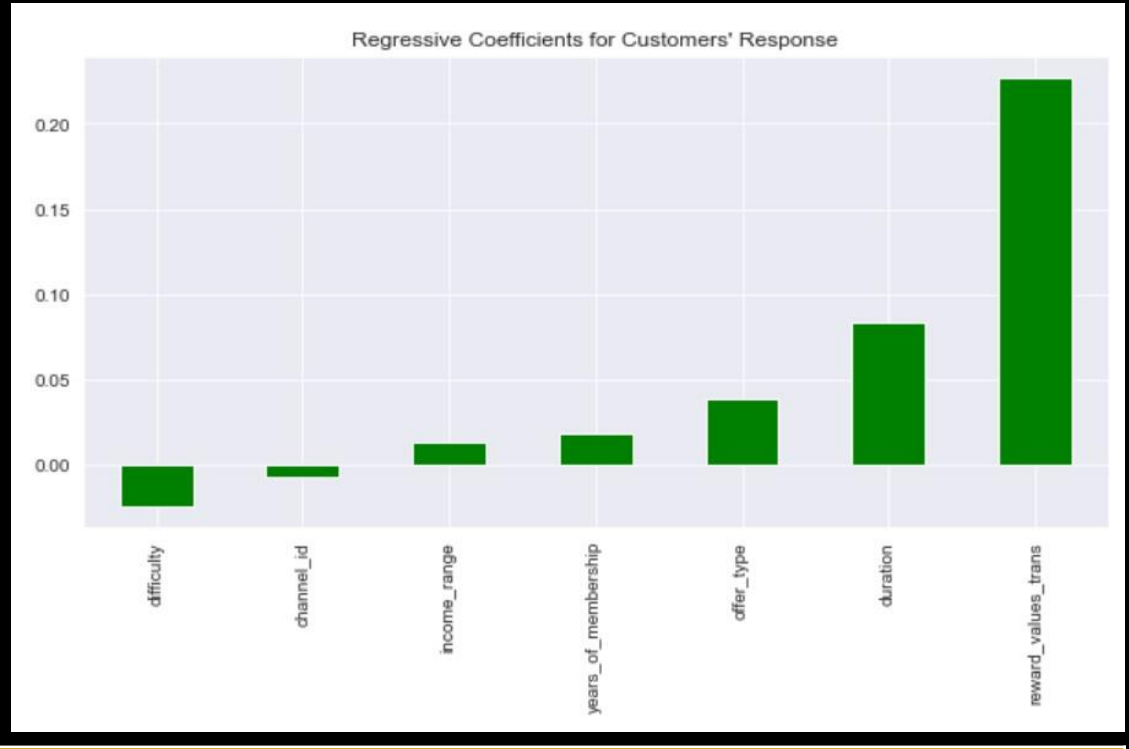
Key Insights

- **MEMBERSHIP: shows steady growth year over year.**
- **Years 2013 and 2018 only have 6 months data and 2018 seems to be the year with the highest membership by >2000 from 2017 followed by year 2017 and 2016.**
- **Offer Type 1 – BOGO provides highest rewards $ value ($5-$10) while Offer Type 2 – DISCOUNT provides lower rewards $ value ($2-$5)**

(3)
How can the Starbucks improve the factors affected Reward members' responses?

Which factors affect members' response and which ones are correlated? Is there room for improvement?





Key Insights

- The correlation matrix didn't show the attributes having very strong correlation. However, the following attributes are correlated: channel_id, difficulty, duration, income_range, reward_value, response and years_of_membership.
- *'reward values'*, *'duration'* and *'offer type'* have strong positive relationship with response followed by *'years of membership'* and members' *'income range'*.
- 'difficulty' is inversely related to response in the regressive model while it showed weak correlation earlier.

Model 1

OLS Regression Results

Dep. Variable:	response	**R-squared:**	0.497
Model:	OLS	**Adj. R-squared:**	0.497
Method:	Least Squares	**F-statistic:**	1.677e+04
Date:	Sun, 06 Jun 2021	**Prob (F-statistic):**	0.00
Time:	15:10:05	**Log-Likelihood:**	-98390.
No. Observations:	118744	**AIC:**	1.968e+05
Df Residuals:	118736	**BIC:**	1.969e+05
Df Model:	7		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	0.0743	0.010	7.365	0.000	0.055	0.094
offer_type	0.0318	0.003	12.591	0.000	0.027	0.037
channel_id	-0.0099	0.002	-4.603	0.000	-0.014	-0.006
difficulty	-0.0241	0.001	-44.109	0.000	-0.025	-0.023
duration	0.0797	0.001	66.591	0.000	0.077	0.082
income_range	0.0094	0.001	6.457	0.000	0.007	0.012
reward_values_trans	0.2263	0.001	328.699	0.000	0.225	0.228
years_of_membership	0.0128	0.001	9.439	0.000	0.010	0.015

Omnibus:	575429.151	**Durbin-Watson:**	1.994
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	13027.692
Skew:	0.110	**Prob(JB):**	0.00
Kurtosis:	1.392	**Cond. No.**	80.8

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Model 2

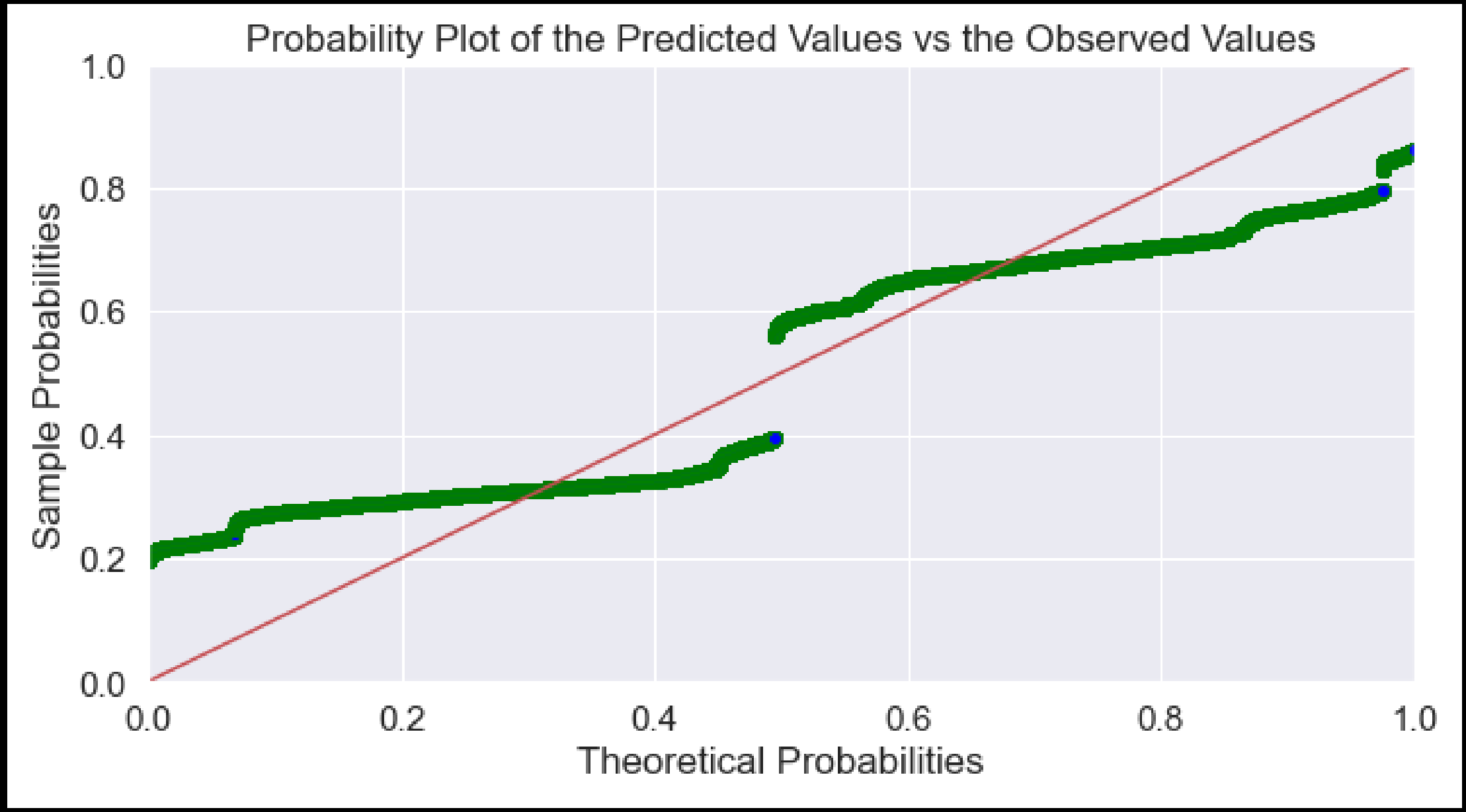
OLS Regression Results

Dep. Variable:	response	**R-squared (uncentered):**	0.744
Model:	OLS	**Adj. R-squared (uncentered):**	0.744
Method:	Least Squares	**F-statistic:**	4.923e+04
Date:	Sun, 06 Jun 2021	**Prob (F-statistic):**	0.00
Time:	15:10:06	**Log-Likelihood:**	-98417.
No. Observations:	118744	**AIC:**	1.968e+05
Df Residuals:	118737	**BIC:**	1.969e+05
Df Model:	7		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
offer_type	0.0388	0.002	16.634	0.000	0.034	0.043
channel_id	-0.0070	0.002	-3.282	0.001	-0.011	-0.003
difficulty	-0.0246	0.001	-45.427	0.000	-0.026	-0.024
duration	0.0831	0.001	75.374	0.000	0.081	0.085
income_range	0.0129	0.001	9.396	0.000	0.010	0.016
reward_values_trans	0.2269	0.001	332.395	0.000	0.226	0.228
years_of_membership	0.0186	0.001	16.642	0.000	0.016	0.021

Omnibus:	579165.172	**Durbin-Watson:**	1.994
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	12936.181
Skew:	0.105	**Prob(JB):**	0.00
Kurtosis:	1.397	**Cond. No.**	20.7

Notes:
[1] R^2 is computed without centering (uncentered) since the model does not contain a constant.
[2] Standard Errors assume that the covariance matrix of the errors is correctly specified.


Probability Plot of the Predicted Values vs the Observed Values
Sample Probabilities
Theoretical Probabilities
0.0
0.2
0.4
0.6
0.8
1.0

The Starbucks Rewards program is very useful both for the company and the customers. Making few adjustments on existing offers that match the members' needs will help the Starbucks to maximize loyalty, increase revenue and satisfy loyal customers by responding to their needs.

Key Takeaways

- Both BOGO and DISCOUNT have similar number of response counts but is Offer Type 1 – BOGO provides the highest rewards $ value ($5-$10) while Offer Type 2 – DISCOUNT provides lower rewards $ value ($2-$5)
- The OLS Regression Analysis created the predictive model that identified the regressive coefficients that are strongly related to members' responses for the following predictive variables :
  - *reward values*
  - *duration*
  - *offer type*
- 74% of the data fit the regression model but there is room for improvement because the differences between the observation and predicted values overlap and break.
- The observed values break may show where values in the model not doing well which leads to a conclusion that another modeling approach can be used to improve the predicted values.

Recommendation

- Investigate why the reward values got highest response for values between $5-$10 and relate future offers with special focus on duration & offer type.
- Include locations of the Starbucks stores that participate in the rewards program to track members' preferences.
- Include datetime stamp in the transcript dataset to track members' ordering & purchasing habits.
- Include the offer ids in the current transcript data under events to differentiate the completed offers by offer type.


End of Presentation

Thank you!