1) **Introduction**

Our group chose a dataset from Starbucks that recorded data on their customers based on gender, age, date of membership, income, how they interact with promotional deals, and how much they spend.

2) **Marketing Questions**

By taking into account the different data that we have, we want to answer our main question "How much do we expect a customer with a Starbucks app account to spend on their next order?" We will further dissect this question by understanding the factors that can influence how much people spend at Starbucks.

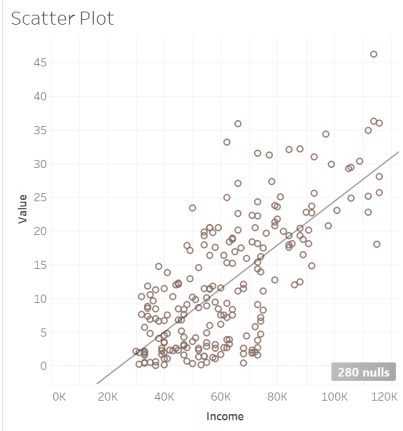
a. Does spending increase as customer income increases?
b. What age demographic tends to spend the most at Starbucks?
c. What gender are the majority of our customers?
d. Does income level affect how often offers are completed?

3) Data Cleanup

person	gender	dummyMale	dummyFemale	age	became_member	days_since_membership	income	event	dummyViewed	dummyCompleted	dummyTransactio	value
d885cede482e4a03b42991c72c254265	F	0	1	29	20170318	1385	61000	transaction	0	0	1	7.51
9b4b02ab6a744385ae6007e4c95099cc	M	1	0	22	20151229	1830	38000	offer received	0	0	0	0
705644ad87c44c09bfc8162c85ec7368	F	0	1	53	20170823	1227	102000	offer received	0	0	0	0
70e5c106ec084acabc0644f12ced1069	F	0	1	59	20150910	1940	35000	offer received	0	0	0	0

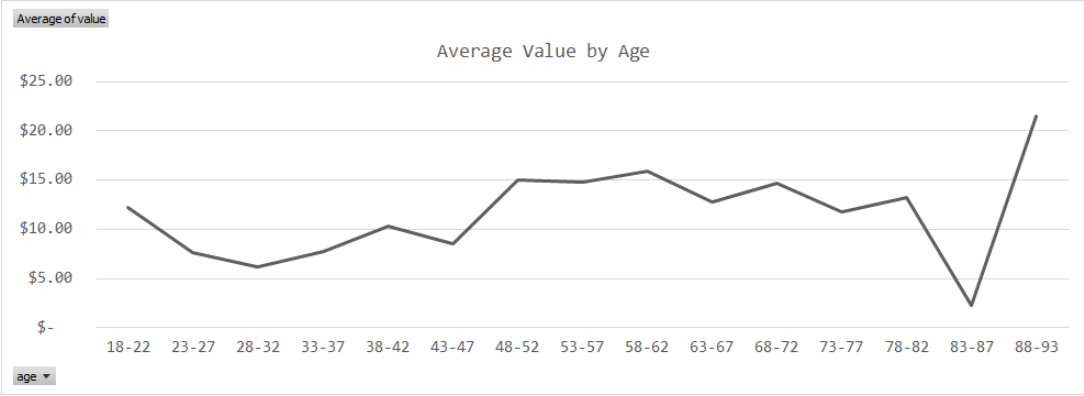
We cleaned the data by first removing data that's duplicated, incomplete, incorrect, irrelevant, or not properly formatted. We then turned categorical data into values by using dummy variables. We separated gender into female and male and event into promotion viewed, completed, or transaction. We also added a column to include days since membership to evaluate how long consumers have been using the Starbucks app.

4) Analysis By Section

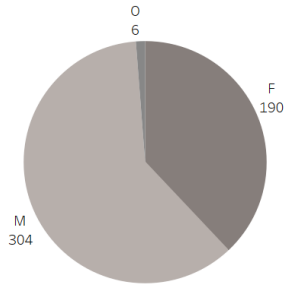


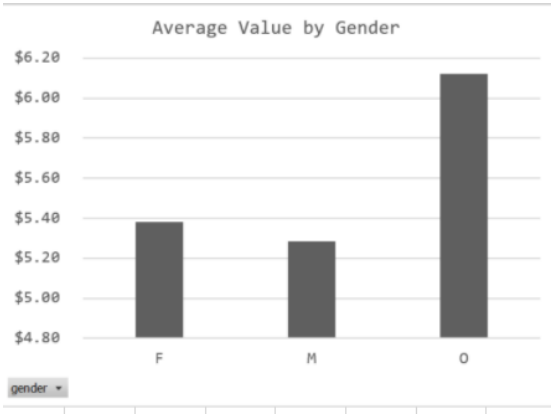
Our first subquestion addresses whether spending increases as customer income increases. We used a scatter plot to visualize the linear regression relationship between value spent in comparison to income.

The linear regression equation is y = -7.91 + .00032(income) showing a positive correlation between income and value spent. The R^2 value of 49% and p-value of <.0001 shows that there is a relationship between income and how much is spent in stores. As expected, those with higher incomes spend more, but the data is fairly spread from the mean. This is likely because Starbucks has low price elasticity due to brand loyalty meaning that members of the Starbucks app are already loyal to the brand so the amount of disposable income they have does not greatly affect their spending habits and a change in higher or lower prices will not affect whether they want to buy Starbucks or not.



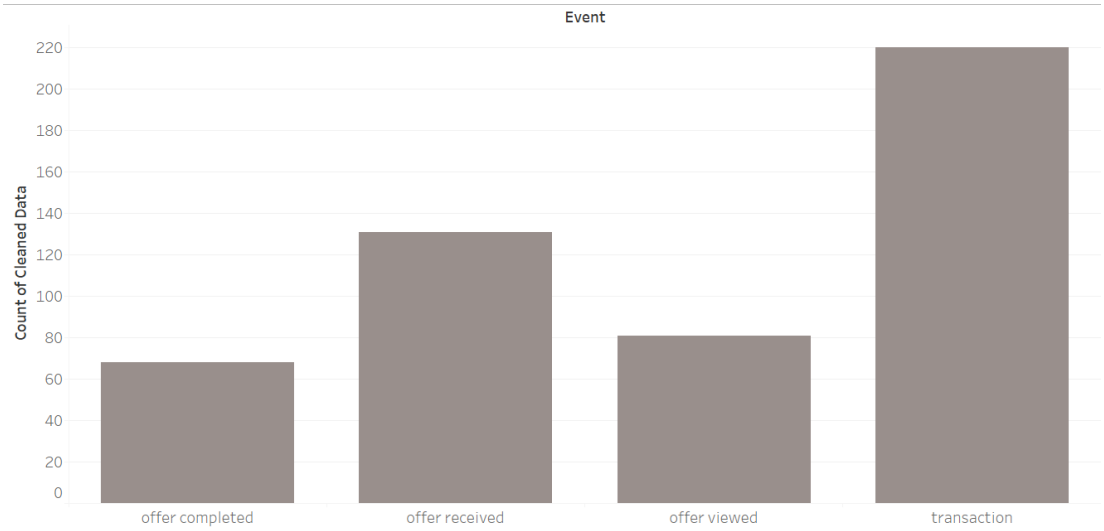
Our second subquestion addresses which age demographic tends to spend the most at Starbucks. We used a line graph to compare the average value spent with age. We found that age 88 spends the most at an average of $35 and age 49 has the second-highest average spend at $31. Ages below 49 generally have a lower spend with 45 being the lowest. It is likely that the older demographic has a stronger affinity towards Starbucks and over the years has built high brand loyalty causing them to spend more at the store. Spending money on gift cards for the holidays for their loved ones may also be a reason for a higher average spend.

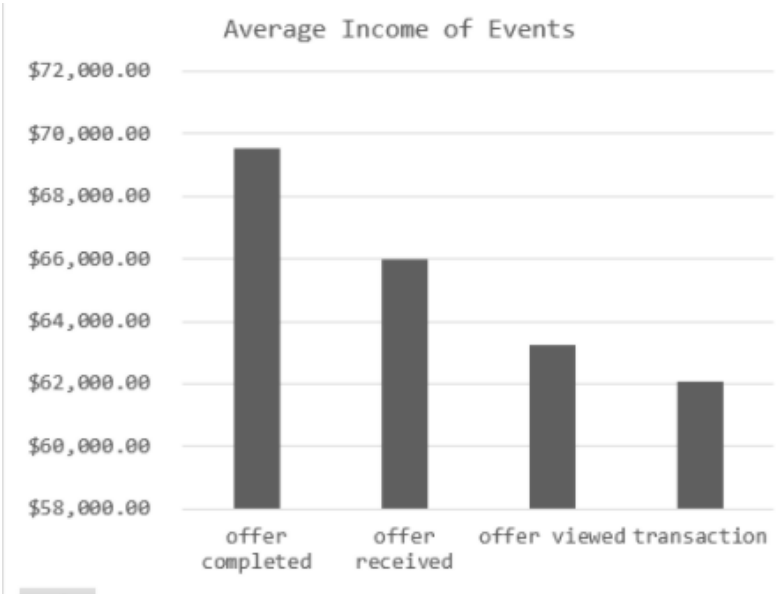




Our third subquestion addresses what gender makes up the majority of Starbucks customers. We created a pie graph comparing the gender breakdown of Starbucks app users. We can see that men make up the majority of Starbucks app users with over 60% of users based on the sample. However, when we created a

bar graph to compare the average value spent based on gender, we found that while men make up the majority of purchases, women on average tend to spend more. This tells us that there is an opportunity for Starbucks to bring in more female customers which can lead to higher profit.

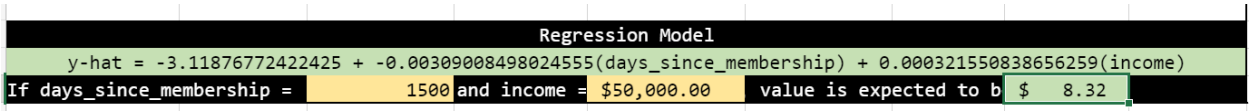




Our fourth subquestion addresses how income level affects how often offers were completed. We first created a bar graph to find out what the most common interaction with a promotional deal is which is a transaction. This indicates that 44% of users tend to make transactions without using an offer. 26% of users received an offer but only 13.6% of users completed an offer meaning that there is room for promotions to be more appealing for more users to click on the offer and convert. We then created a bar graph to compare the events to income level. We found that people with lower income tend to complete transactions while people with higher income tend to complete offers. It is likely that offers require a

certain level of spending to be completed which is why people with lower income are not taking advantage of it.

5) Conclusions

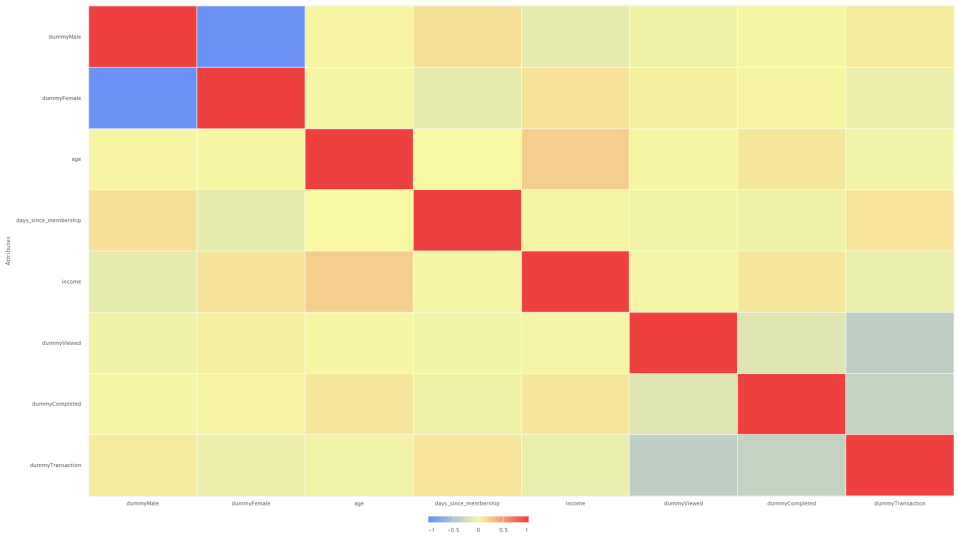


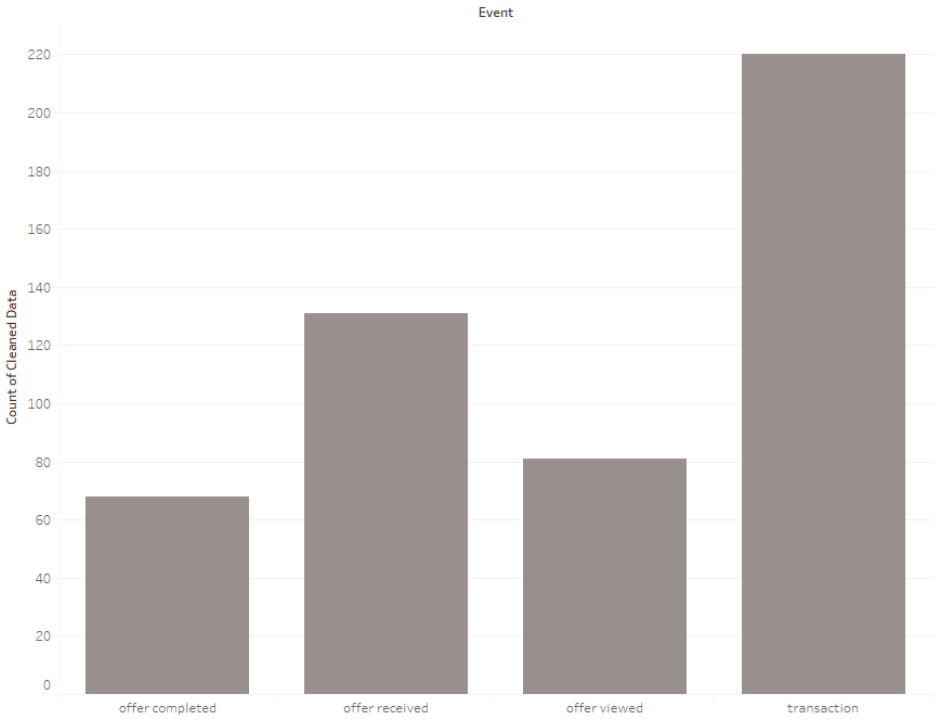
To answer our main question of how much we expect a customer to spend, we ran a regression analysis, first using all variables, and then cutting it down to only the variables with significant p-values: income and days since membership. We found a negative relationship between value spend and days since membership while income and value spend have a positive relationship. Plugging in values of days since membership and income give us an approximation of how much they will spend with an R^2 value of 52%. Through our questions we were able to understand that income and length of membership are related to customer spending, ages 49-88 tend to spend more at Starbucks, men make up the majority of app users but tend to spend less than women, and those with higher income tend to take advantage of promotions more.

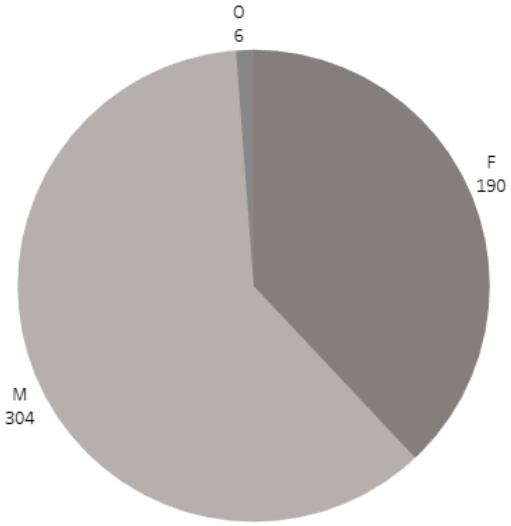
6) Recommendations

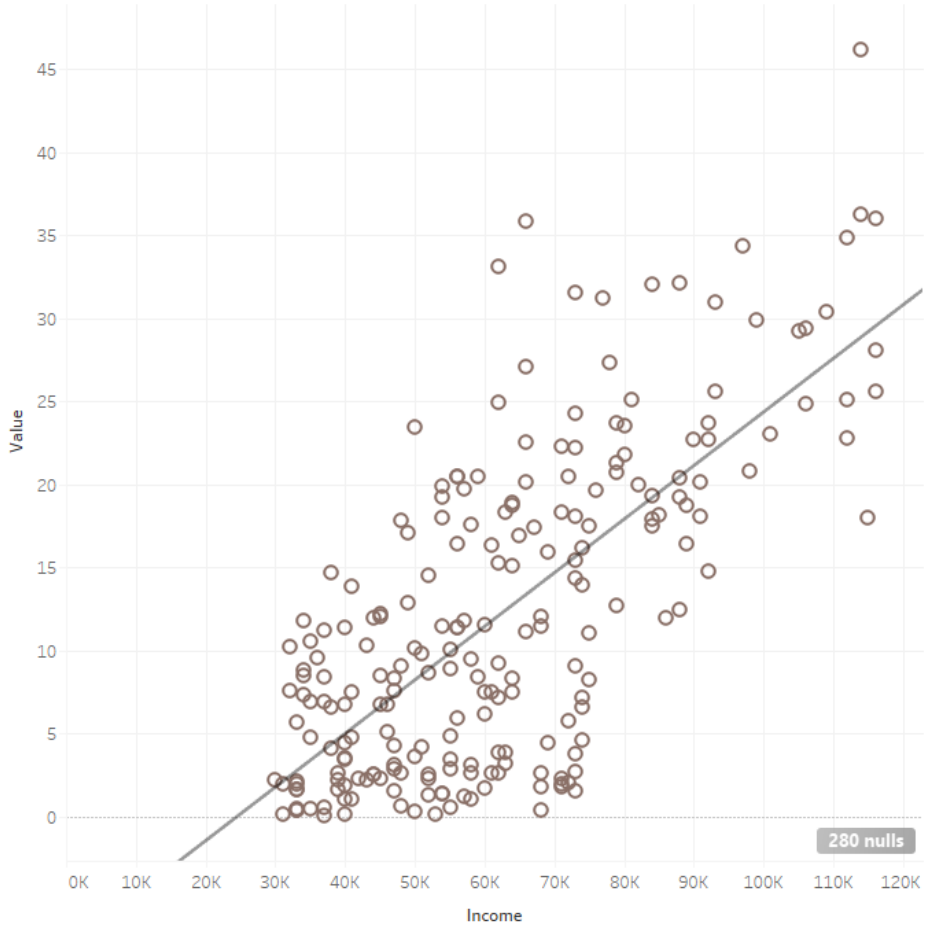
As a marketer, I recommend targeting an age demographic of 49+, especially during the holidays to promote gift card sales as we've seen a high spend from this demographic that's way above the average price of a drink. This is possible through spending money on gift cards which makes sense for marketers to promote this messaging. Another recommendation is to target women promoting them to become Starbucks app users because on average they tend to spend more than men. This is great to help increase revenue for Starbucks. Our last recommendation is to make promotion deals easier to complete so that those with lower income can take advantage of the promotions as well. Starbucks' app should be more inclusive of those from different income levels and motivate them to convert.

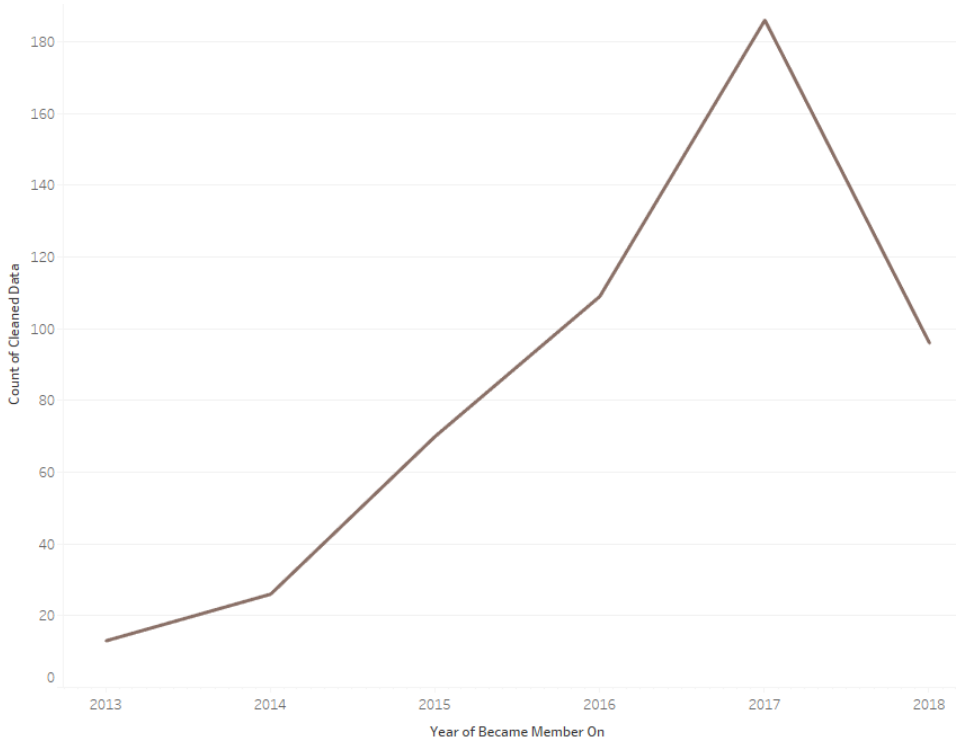
7) Appendix - Screenshots

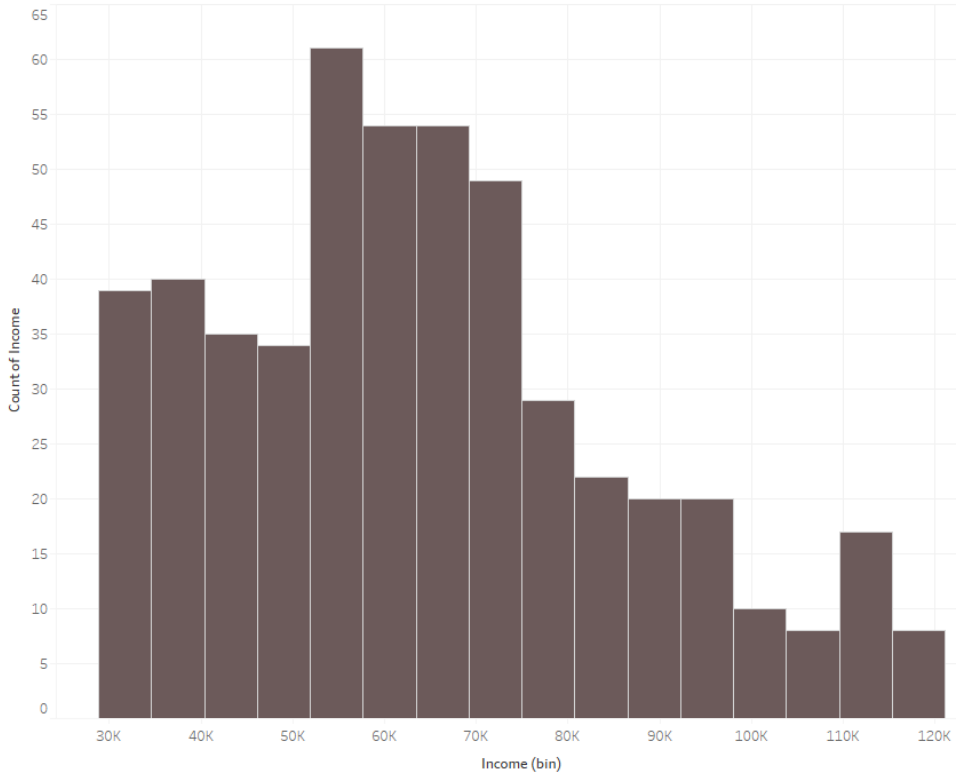





O
6
F
190
M
304


45
40
35
30
25
20
15
10
5
0
Value
280 nulls
0K
10K
20K
30K
40K
50K
60K
70K
80K
90K
100K
110K
120K
Income


Count of Cleaned Data
180
160
140
120
100
80
60
40
20
0
2013
2014
2015
2016
2017
2018
Year of Became Member On


Count of Income
65
60
55
50
45
40
35
30
25
20
15
10
5
0
30K
40K
50K
60K
70K
80K
90K
100K
110K
120K
Income (bin)

Univariate Statistics

	gender	age	became_member_on	days_since_membership	income	event	value
Mean	N/A	54	N/A	1479	$ 64,304.00	N/A	$ 5.33
Median	N/A	55	N/A	1354	$ 62,000.00	N/A	$ -
Mode	M	57	2015-09-10	1940	$ 55,000.00	transaction	$ -

Predicting Value

Regression Statistics	
Multiple R	0.719275552
R Square	0.51735732
Adjusted R Square	0.512909
Standard Error	6.762528798
Observations	220

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	10637.58059	5318.790293	116.3039895	4.71402E-35
Residual	217	9923.799677	45.73179575		
Total	219	20561.38026			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	-3.118767724	2.135629557	-1.460350515	0.001456401	-7.328000255	1.090464807	-7.328000255	1.090464807
days_since_members	-0.003090085	0.00103019	-2.999527786	0.003019863	-0.005120545	-0.00105962	-0.005120545	-0.00105962
income	0.000321551	2.15858E-05	14.89640918	4.87851E-35	0.000279006	0.000364095	0.000279006	0.000364095

Regression Model

y-hat = -3.11876772422425 + -0.00309008498024555(days_since_membership) + 0.000321550838656259(income)

Data Dictionary	
person	Unique identifier for each observation
gender	Gender of user; M, F, or O
dummyMale	Dummy variable; 1 if M, 0 if not
dummyFemale	Dummy variable; 1 if F, 0 if not
age	Age of user; years
became_member_on	Date user registered Starbucks app account; YYYYMMDD
days_since_membership	Calculated variable; difference in days between 2021-01-01 and became_member_on
income	Annual income of user; US dollars
event	Action taken by user; offer received, offer viewed, offer completed, or transaction
dummyViewed	Dummy variable; 1 if offer viewed, 0 if not
dummyCompleted	Dummy variable; 1 if offer completed, 0 if not
dummyTransaction	Dummy variable; 1 if transaction, 0 if not
value	Value of transaction, if any; US dollars