**Offers.. Offers.. Offers!**

**From Starbucks**

A quick analysis review of simulated data provided by Starbucks to identify which groups of people are most responsive to each type of offer sent by Starbucks app.

****

[VIA @STARBUCKS/INSTAGRAM](https://www.instagram.com/p/CL6zqJOsypO/)

1. **Introduction**

Once every few days, Starbucks sends out an offer to users of the mobile app. An offer can be merely an advertisement for a drink or an actual offer such as a discount or BOGO (buy one get one free)

A. In a BOGO offer, a user needs to spend a certain amount to get a reward equal to that threshold amount.

B. In a discount, a user gains a reward equal to a fraction of the amount spent.

C. In an informational offer, there is no reward, but **neither** is there a requisite amount that the user is expected to spend. Offers can be delivered via multiple channels.

1. **Problem Statement**

The basic task is to use the data to identify which groups of people are most responsive to each type of offer.

For more detailed information, you can check the full project on GitHub repository here:

<https://github.com/michael-fawzy/DSND-Capstone>

1. **Data**

The data set used contains simulated data that mimics customer behavior on the Starbucks rewards mobile app. The program used to create the data simulates how people make purchasing decisions and how those decisions are influenced by promotional offers.

The data is contained in three files that can be found in the project GitHub repository mentioned above:

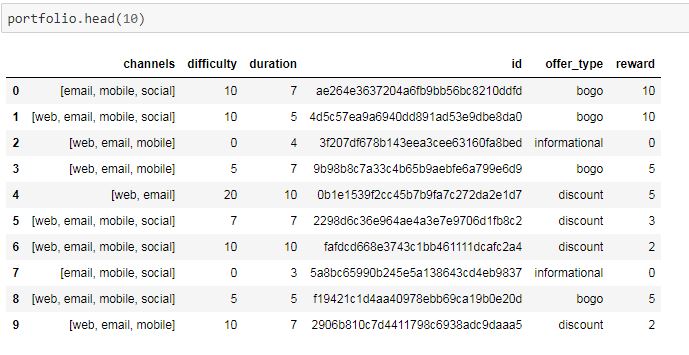
* portfolio.json - containing offer ids and meta data about each offer (duration, type, etc.)
* profile.json - demographic data for each customer
* transcript.json - records for transactions, offers received, offers viewed, and offers completed

Here is the schema and explanation of each variable in the files:

***A. Portfolio.json***

Offers sent during 30-day test period (10 offers x 6 fields)

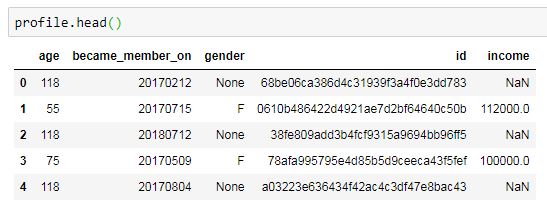
* id (string) - offer id
* offer\_type (string) - type of offer ie BOGO, discount, informational (an advertisement for a drink).
* difficulty (int) - minimum money required to be spent to complete an offer and receive reward.
* reward (int) - reward given for completing an offer (money awarded for the amount spent).
* duration (int) - time for offer to be open, in days.
* channels (list of strings) - web, email, mobile, social.



***B. Profile.json***

Rewards program users (17000 users x 5 fields)

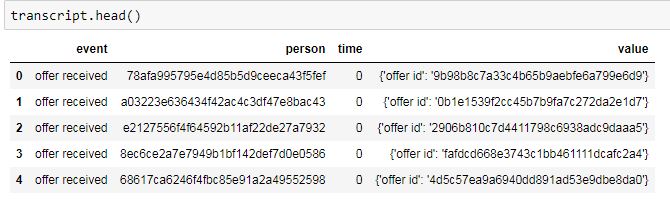
* age (int) - age of the customer. Missing value encoded as 118
* became\_member\_on (int) - date when customer created an app account (date format YYYYMMDD)
* gender (str) - gender of the customer (note some entries contain 'O' for other rather than M or F)
* id (str) - customer id
* income (float) - customer's income



***C. Transcript.json***

Event log (306648 events x 4 fields)

* event (str) - record description (ie transaction, offer received, offer viewed, offer completed).
* person (str) - customer id
* time (int) - time in hours since start of test. The data begins at time t=0
* value - (dict of strings) - different values depending on event type:
  + offer id: (string/hash) not associated with any "transaction"
  + amount: (numeric) money spent in "transaction"
  + reward: (numeric) money gained from "offer completed"

****

1. **Strategy**

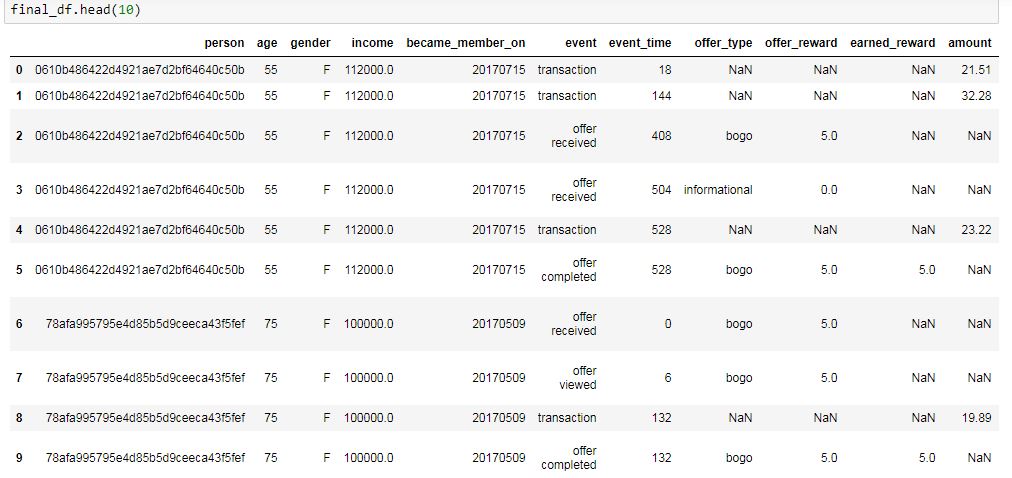
I have decided to develop a set of heuristics to help determine what offer should be sent to each customer through the following steps:

1- Combine transaction, demographic and offer data after necessary cleaning.

2- Find correlation between income and purchase.

3- Determine which demographic group pays more in purchases.

4- Determine which demographic groups respond best to which offer type.



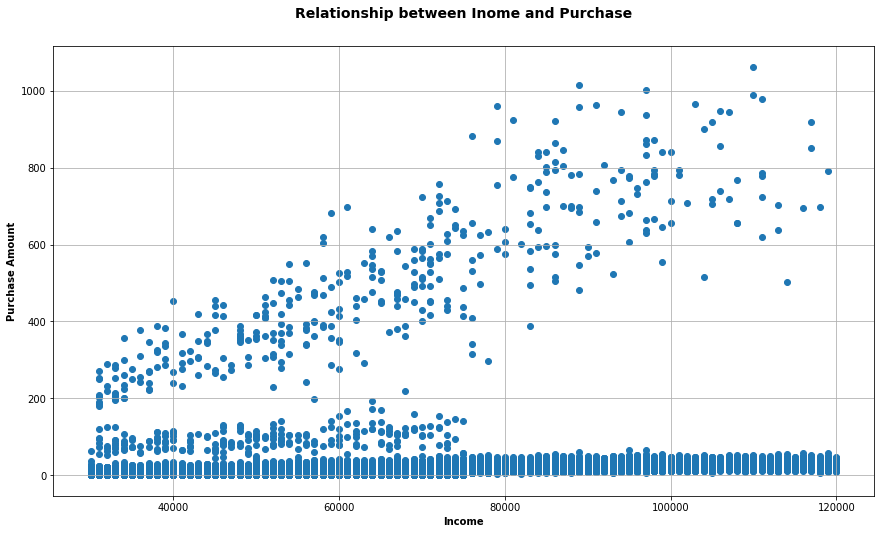
1. **Questions & Answers**

***Q1- Is there a correlation between income and purchase amount?***

Although there is a direct correlation between customer income and their purchase amount, it seems there is a group of customers whose purchase is constant whatever their income is. Hmm.. Interesting!

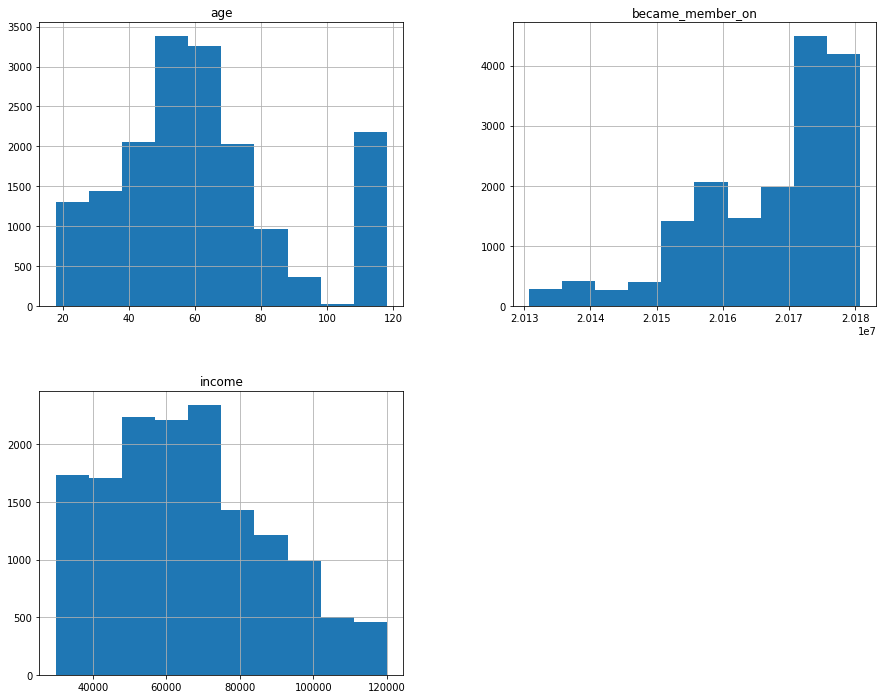
I believe Starbucks need to dig deeper into this to find what traits those customers have.

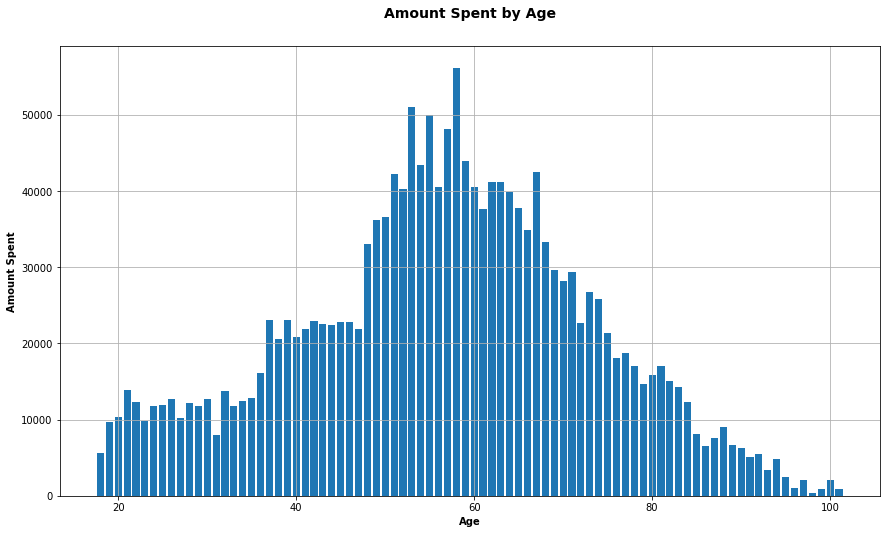
Meanwhile, the company should target high income customers as they are willing to pay more.

****

***Q2- Do purchases correspond to age distribution?***

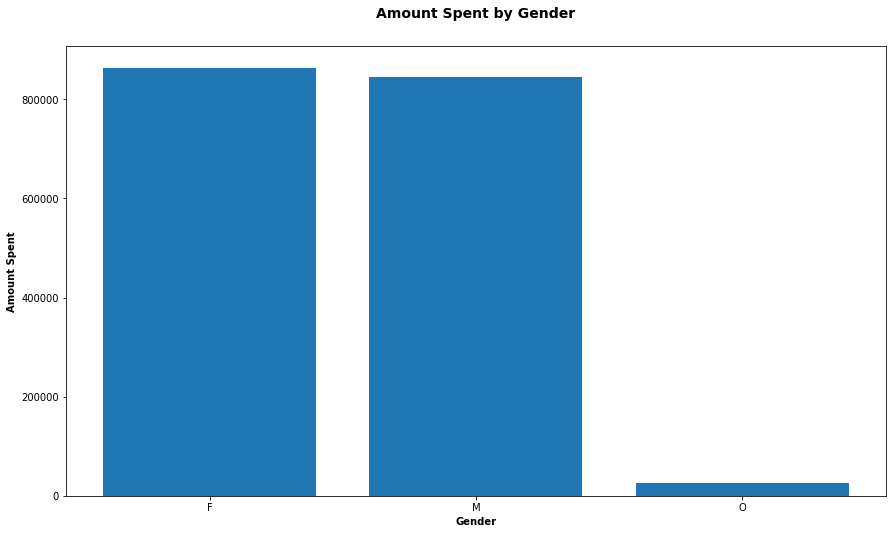
Yes. The amount spent corresponds to age distribution. Customers between 50 & 70 years spend the most. I believe this indicates regular spending rate among all customers. In other words, customers of all age groups spend the same average value, thus the more the customers, the higher the amount they spend. So, the company should focus more on customer groups of higher member counts.

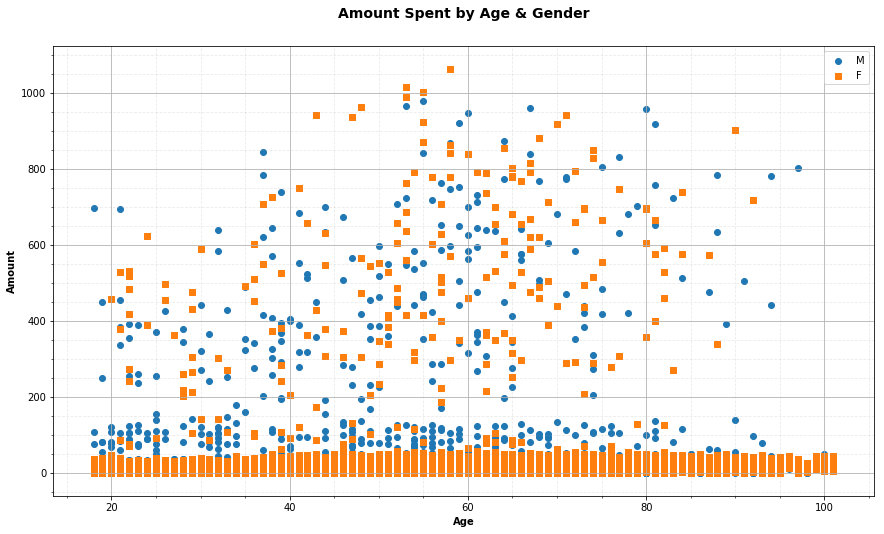
****



***Q3- Which demographic groups pays more in purchases?***

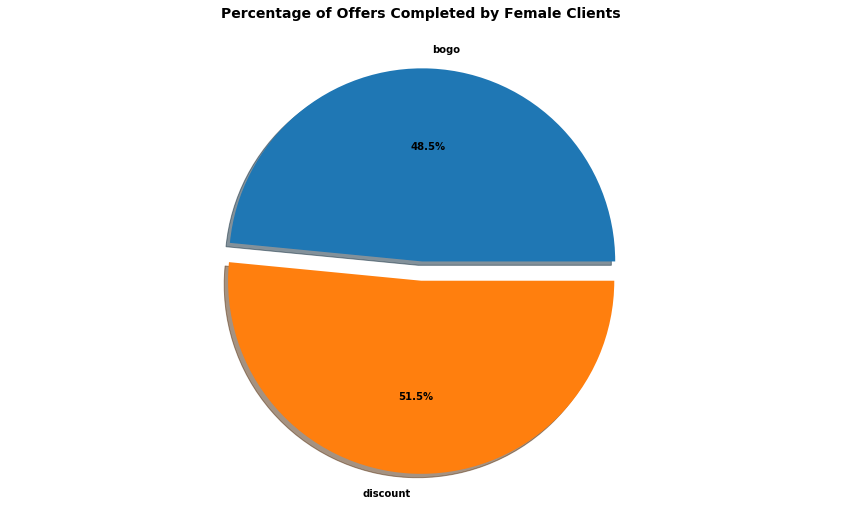
In general, females spend the most (863695), but are very close to males (844890)

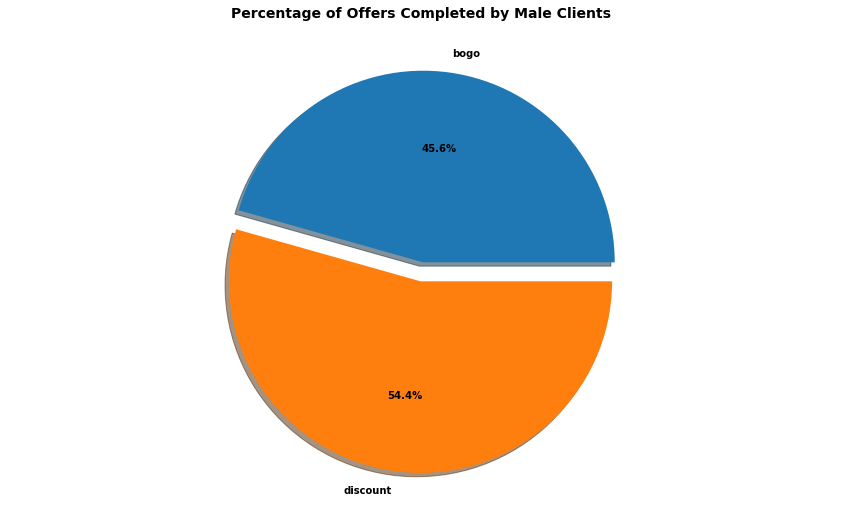




According to age, females of age 43-58 spend the most while males of age 53-80 spend the most.

***Q4- Which demographic groups respond best to which offer type?***

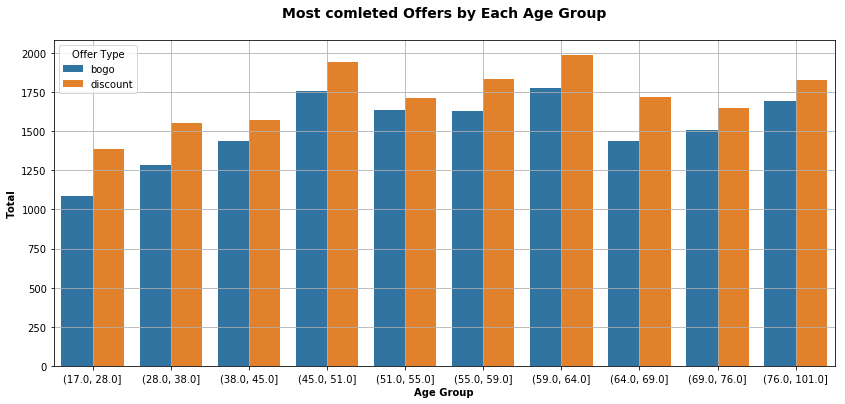




Both males and females responded more to discount offers.

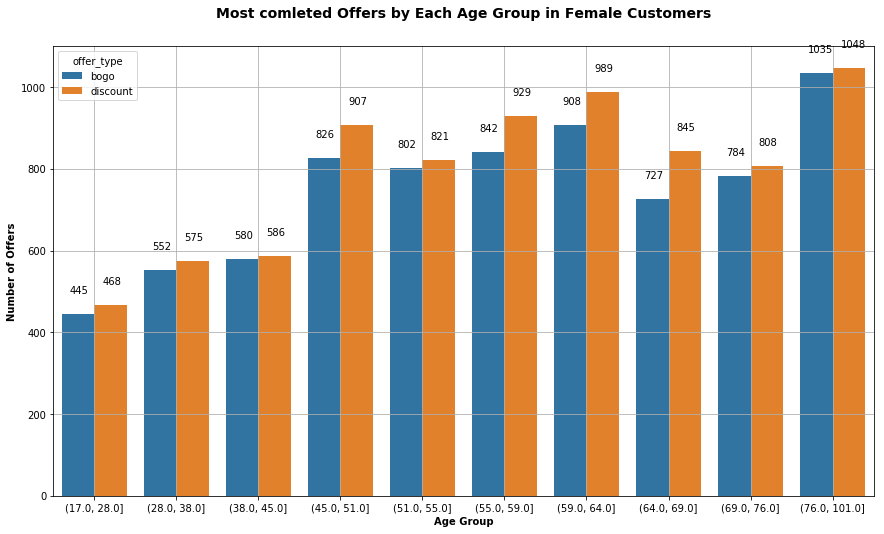
More males (54.4%) responded to discount offers than females (51.5%).

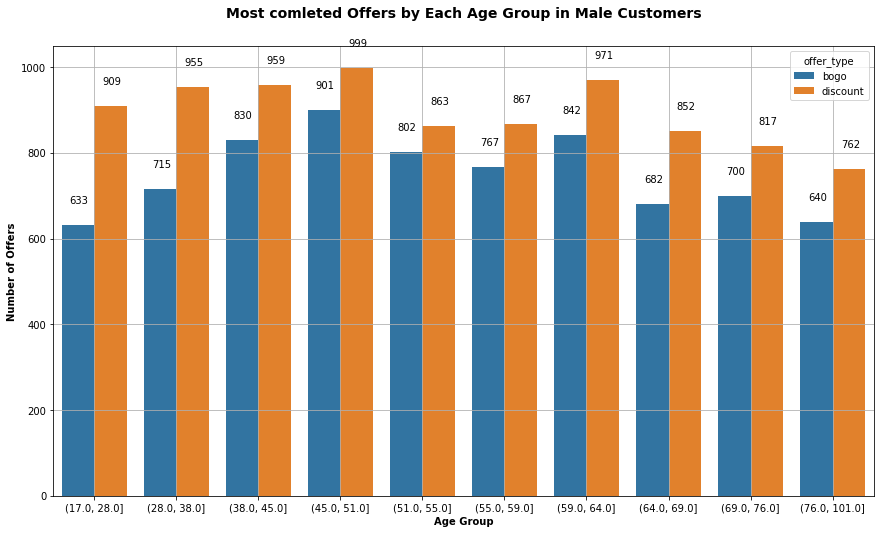
More females (48.5%) responded to bogo offers than males (45.6%).



Customers of all age groups tend to complete discount offers more than bogo offers.

Collectively, customers of age 59-64 tend to complete offers the most.





When viewed separately, the highest number of offers was completed by male customers of age 45-51, followed by those of age 59-64 while the female customers of age 76-101 completed the highest number of offers, followed by those of age 59-64.

1. **Final conclusions**

According to the problem statement, the basic task is to use the data to identify which groups of people are most responsive to each type of offer. After this analysis, I can now say that all customers are more responsive to discount offers than they are to bogo offers although more females (48.5%) responded to bogo offers than males (45.6%).

Since there is a direct correlation between customer income and the amount they spend, we can safely assume that Starbucks should focus more on targeting customers of higher income as they tend to pay more.

An interesting fact, however, that requires more research, is why there is a group of people whose purchase is not directly correlated to their income and keep a fixed payment amount.

On a side note, it was also interesting to find that some customers were 100 years old, and even more, who not only still make purchases from Starbucks, but they also use a mobile app. to do this.! Wow! Now that's what I call adaptation to technology..

**As a final conclusive result, the company should focus more on sending discount offers to higher income customers of all genders in the age group 59-64 since this group tends to pay more and complete more offers. However the data still hides a lot under the hat, and the possibilities look endless!**

Link:

<https://michaelfawzy.medium.com/offers-offers-offers-from-starbucks-88707094a5fa>