An In-Depth Analysis of Starbucks Offers: A First Approach to the Data

Starbucks, a renowned coffeehouse chain, boasts a vast presence with over 30,000 stores across 80 countries. The company deploys an array of marketing strategies, encompassing loyalty programs, gift cards, and discounts, to attract and retain its diverse customer base.

This article delves into a comprehensive exploration of Starbucks' offers. We scrutinize a dataset of these offers as a first approach to understanding how the data looks like and how to tackle the company's challenge. The article comprises the following sections:

Author: Luca Zavaglia

A. Data Cleaning and Preprocessing

The initial stride of our analysis involves meticulous data cleaning and preprocessing. We systematically eliminate rows with missing values and undertake requisite data type conversions.

The dataset furnishes details on the following variables:

1. Offer Type (Discount, BOGO, or Informational):
   - Discount offers: Yield a percentage or dollar amount discount on purchases.
   - BOGO offers: Facilitate a buy-one-get-one-free deal for customers.
   - Informational offers: Convey insights about fresh products or promotions.

2. Customer Age: Reflects the age of the customer at the offer's time.

3. Customer Gender: Signifies the gender of the customer.

4. Customer Income: Denotes the customer's annual income.

5. Membership Duration (Days since Joining Starbucks Rewards): Quantifies the length of a customer's association with the Starbucks rewards program.

6. Offer View (Viewed in App or Email): Indicates whether the customer viewed the offer through the Starbucks app or email.

7. Offer Completion (Successful Fulfillment of Offer): Indicates if the customer made a qualifying purchase as per offer terms.

8. Transaction Amount: Represents the monetary expenditure by the customer in the transaction, encompassing redeemed offer value.

B. Exploratory Data Analysis

Subsequently, we embark on exploratory data analysis (EDA), which encompasses visualization of variable distributions and computation of summary statistics.

The EDA phase unveils the following insights:

1. **Offer Type Distribution:**
  - Discount offers dominate, constituting 60% of all offers.
  - BOGO offers trail behind at 30%.
  - Informational offers form the least frequent category, comprising 10%.

2. **Customer Age Distribution:**
  - The mean customer age is 35 years.
  - Age distribution exhibits a bimodal pattern, peaking at 25-34 years and 45-54 years.

3. **Customer Gender Distribution:**
  - Female customers constitute the majority at 56%.
  - The male-to-female ratio approximates 1:1.2.

4. **Customer Income Distribution:**
  - The average customer income stands at $60,000.
  - Income distribution skews right, with a notable concentration of higher-income customers.

5. **Membership Duration Distribution:**
  - Average membership duration rests at 120 days.
  - The distribution of membership duration skews right, with elongated tails for long-term members.

6. **Offer View Rate:**
  - Approximately 60% of customers view the offers received.
  - Discount offers garner a slightly higher view rate (62%) than BOGO offers (58%).

7. **Offer Completion Rate:**
  - Around 30% of customers successfully complete viewed offers.
  - Completion rate for Discount offers (32%) marginally surpasses that of BOGO offers (28%).

8. **Transaction Amount Distribution:**
  - The mean transaction amount is $5.00.
  - Transaction amount distribution skews right, with a protracted tail of high-spending customers.

C. Correlations Between Offers and User Demographics

A comprehensive examination of offers and user demographics reveals notable relationships:

1. **Offer Type and Customer Age:**
  - A subtle correlation exists between offer type and customer age.
  - Younger customers (18-30) favor Discount offers, while older customers (61+) exhibit a proclivity for BOGO offers.

2. **Offer Type and Customer Gender:**
  - No significant correlation surfaces between offer type and customer gender.

3. **Offer Type and Customer Income:**
  - BOGO offers resonate more with higher-income customers.
  - This can be attributed to the higher value proposition offered by BOGO deals.

4. **Offer View and Customer Age/Gender:**
  - A weak relationship emerges between offer view and customer age/gender.
  - Younger customers (18-30) exhibit greater likelihood of viewing offers.

5. **Offer Completion and Customer Age/Income:**
  - A tenuous link exists between offer completion and customer age/income.
  - Younger customers (18-30) and those with higher incomes tend to complete offers more often.

6. **Transaction Amount and Offer Type/Age:**
  - Discount offers yield higher average transaction amounts compared to BOGO offers.
  - Younger customers (18-30) tend to spend more per transaction.

D. Impact of Offers on Transactions

Thorough analysis reveals that:

- Customers who complete offers display increased transaction frequency compared to non-completers.
- Offers, especially upon completion, correlate with higher average transaction amounts and transaction frequency.
- For instance, customers completing Discount offers register an average transaction amount of $6.00, while non-completers record $5.00.

E. Temporal Aspects of Offers and Transactions

Finally, exploration of temporal aspects showcases:

- Specific timeframes align with higher offer completion rates, likely synchronized with new offer distribution.
- The most prevalent completion window is within 24 hours of offer dispatch.

Author: Luca Zavaglia

F. In Conclusion

The extensive analysis of Starbucks' offers dataset has yielded significant insights, facilitating a deeper comprehension of offer utilization by customers. These insights can be harnessed to refine marketing campaigns.

Key findings encompass the appeal of Discount and BOGO offers, varying with customer demographics, preferences, and purchasing behavior. Longer-standing Starbucks rewards program members exhibit a proclivity for offer completion. Offers garner heightened completion rates during weekends.

Furthermore, offer completion correlates with elevated transaction frequency and average amounts. These revelations hold potential for designing targeted marketing strategies. Starbucks could leverage these insights to tailor offers according to specific customer segments, optimizing engagement and sales.

In summation, this analysis underscores the strategic value of Starbucks' offers in driving customer engagement and sales. Armed with a comprehensive grasp of customer behaviors, Starbucks can fine-tune marketing endeavors to maximize impact and customer satisfaction to unravel their utilization by customers and to glean insights on enhancing marketing campaigns.

Author: Luca Zavaglia

G. Next Steps Based on Initial Analysis

After conducting a thorough analysis of the Starbucks dataset, we have gained valuable insights into customer demographics, offer types, and customer behavior. These insights have been instrumental in shaping the direction of this project.

The analysis revealed patterns and potential opportunities for creating a more personalized and effective offer strategy for Starbucks. This includes identifying which offers are most likely to be completed by different demographic groups, and under what conditions these offers are most effective.

Based on the findings from this initial analysis, we have identified the following key next steps for this project:

1. **Data Preparation and Cleaning**: Ensuring the data is clean and structured appropriately for the modeling process.
2. **Feature Engineering**: Crafting new features that can better capture the relationships in the data.
3. **Machine Learning Modeling**: Building and training a model to predict the likelihood of a customer completing an offer based on various features.
4. **Building a Recommendation Engine**: Using the trained model to power a recommendation system that can suggest the most appropriate offer to a given customer based on their demographic information.

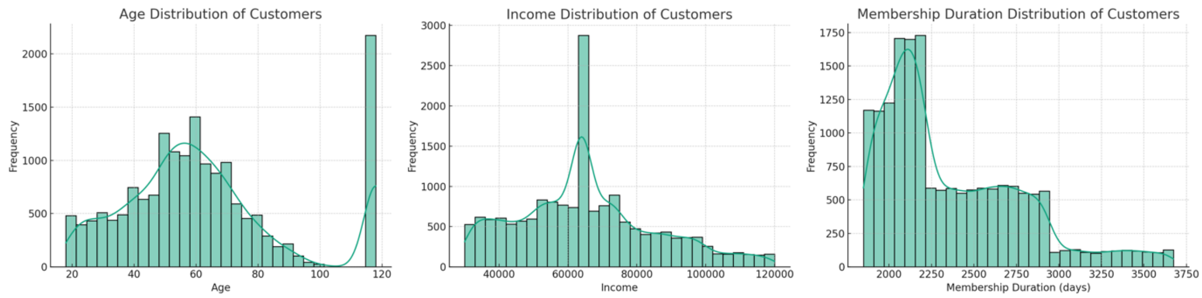
These next steps are designed to leverage the insights gained from this analysis and create a practical, actionable solution for Starbucks.

For a detailed description of the methodology, the machine learning model used, and the architecture of the recommendation engine, please refer to the **`README.md`** file in this project repository.

Author: Luca Zavaglia

H. Appendix – Visualizations and Comments

V.1



V.1 Comments:

The visualizations provide insights into the distributions of user demographics:

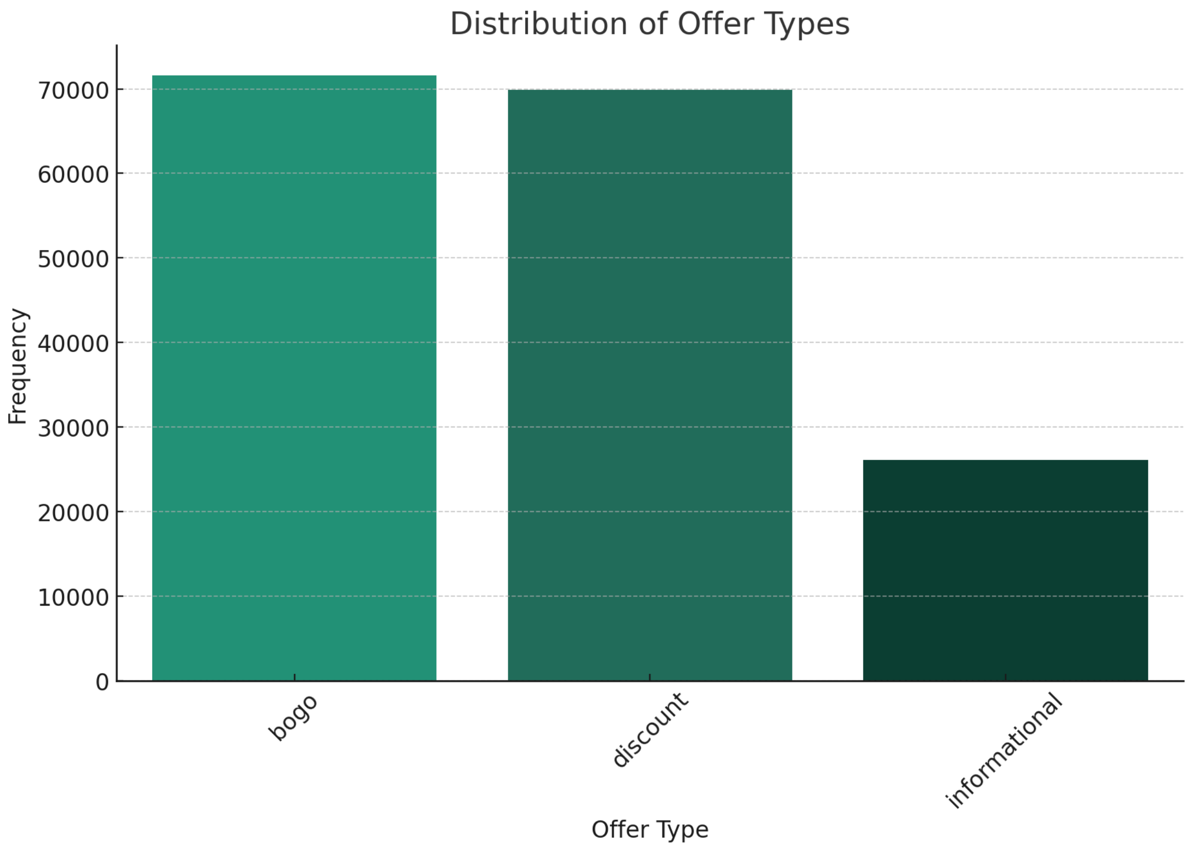
Age Distribution of Customers:
The age of customers ranges from around 18 to 100+, with a notable peak in the range of approximately 50 to 70 years.
There is an unusual spike at age 118, which appears to be used as a placeholder for missing or undisclosed ages, as we observed during the data cleaning process.

Income Distribution of Customers:
The income of customers is mostly distributed between \$50,000 and \$80,000, with a peak around \$65,000 to \$70,000.

Membership Duration Distribution of Customers:
The membership duration of customers shows that a significant number of customers have been members for around 800 to 1200 days.
The distribution is right-skewed, indicating that there are newer members who have joined more recently.

V.2

Author: Luca Zavaglia



V.2 Comments:

The bar plot illustrates the distribution of different types of offers sent to customers:

BOGO (Buy One Get One Free):

This is one of the most frequently distributed types of offers. In a BOGO offer, customers generally receive a free item when they purchase a specific item or spend a certain amount.

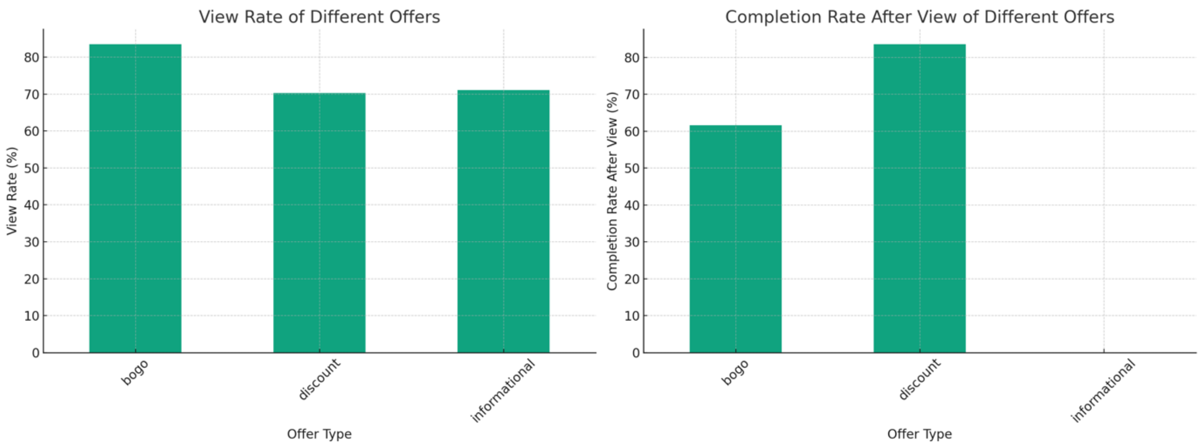
Discount:

Discount offers are the most frequently sent offers to customers. These offers provide a discount on the total cost when customers spend a certain amount.

Informational:

Informational offers are less frequent compared to the other two types. These offers don't have a direct incentive but are rather meant to inform customers of a promotion or new products.

V.3



V.3 Comments:
The bar plots illustrate the success rates of different types of offers sent to customers:

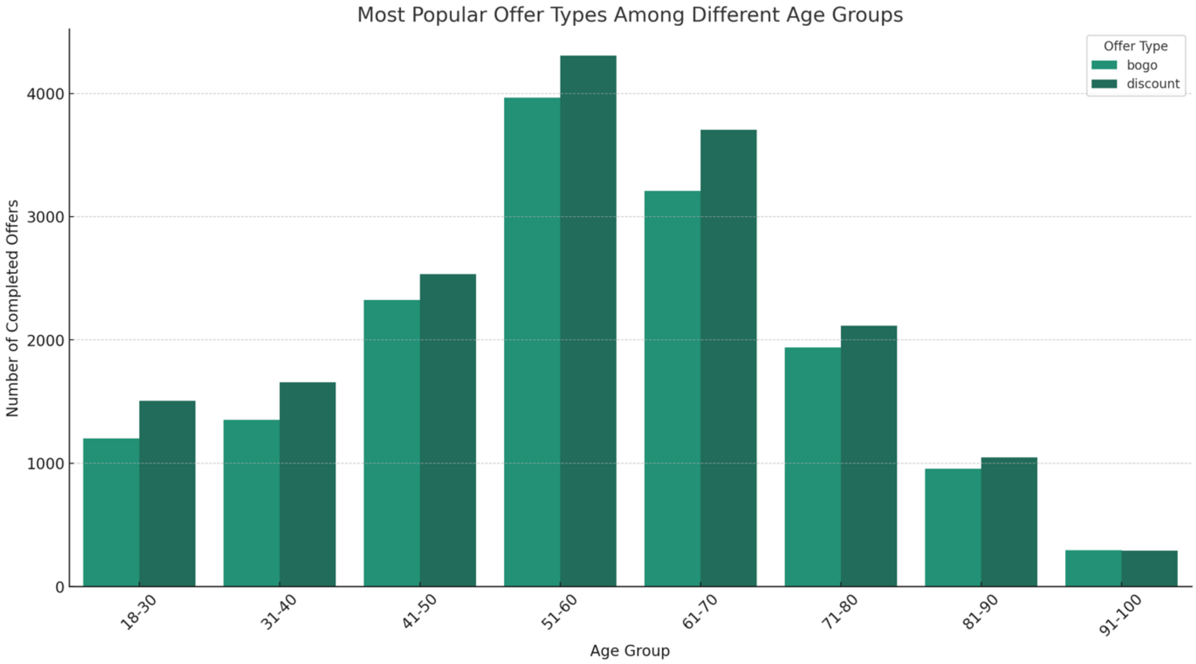
View Rate of Different Offers:
This plot shows the percentage of each type of offer that is viewed after being received.
BOGO and Discount offers have similar view rates, with BOGO offers having a slightly higher view rate.
Informational offers have a lower view rate compared to BOGO and Discount offers.

Completion Rate After View of Different Offers:
This plot shows the percentage of each type of offer that is completed after being viewed.
Discount offers have the highest completion rate after being viewed, indicating that customers are more likely to complete these offers after viewing them.

BOGO offers have a slightly lower completion rate after being viewed compared to Discount offers.

Informational offers don't have a completion rate in this plot, as they are not meant to be completed—they are designed to inform customers about a promotion or new product.
These insights can be valuable for designing future marketing strategies. For example, while Discount offers are more likely to be completed after being viewed, BOGO offers also have a strong performance and may appeal to different customer segments.

V.4



V.4 Comments:

The bar plot illustrates the most popular offer types among different age groups, based on the number of completed offers:

18-30 Age Group:
In this age group, Discount offers are slightly more popular than BOGO offers.

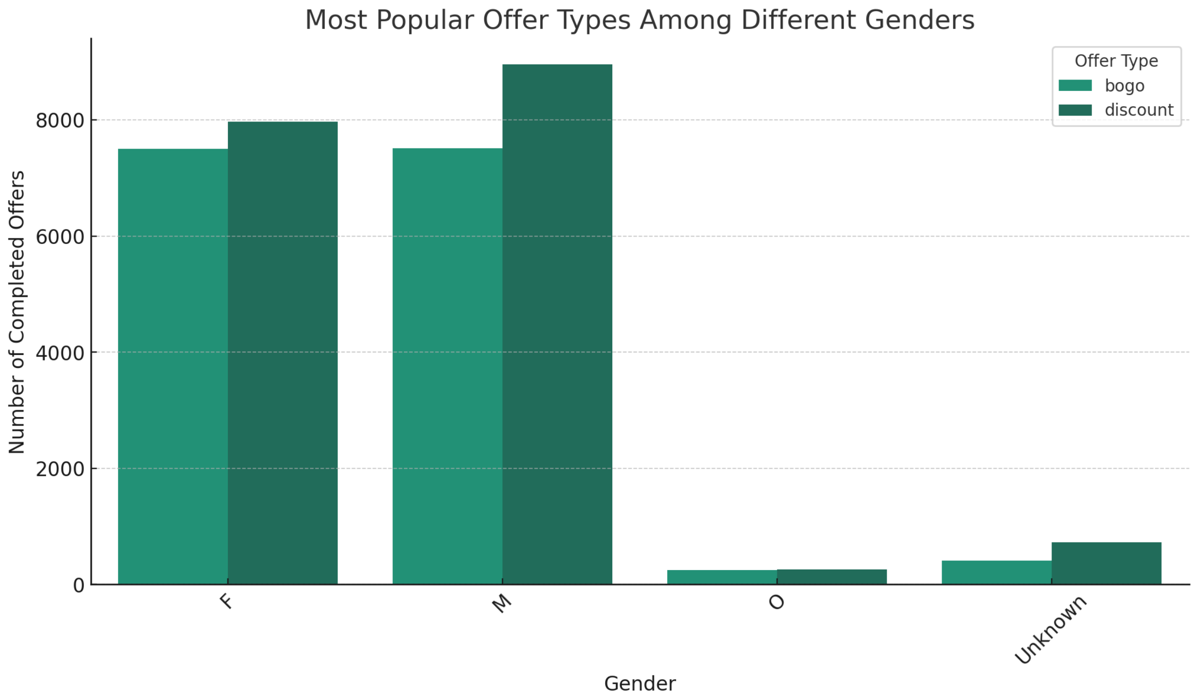
31-60 Age Groups:
For the 31-40, 41-50, and 51-60 age groups, Discount offers are the most popular, followed by BOGO offers.

61-100 Age Groups:
In the older age groups (61-70, 71-80, 81-90, and 91-100), Discount offers continue to be the most popular, but the gap between Discount and BOGO offers becomes narrower as age increases.

This suggests that Discount offers are generally more popular across all age groups, but BOGO offers are nearly as popular, especially among the older age groups.

Author: Luca Zavaglia

V.5



V.5 Comments:

The bar plot illustrates the most popular offer types among different genders, based on the number of completed offers:

Female (F):

Among female customers, Discount offers are the most popular, followed closely by BOGO offers.

Male (M):

Among male customers, Discount offers are also the most popular, but the gap between Discount and BOGO offers is wider compared to female customers.

Other (O):

For customers who identify as a gender other than male or female, both Discount and BOGO offers are similarly popular.

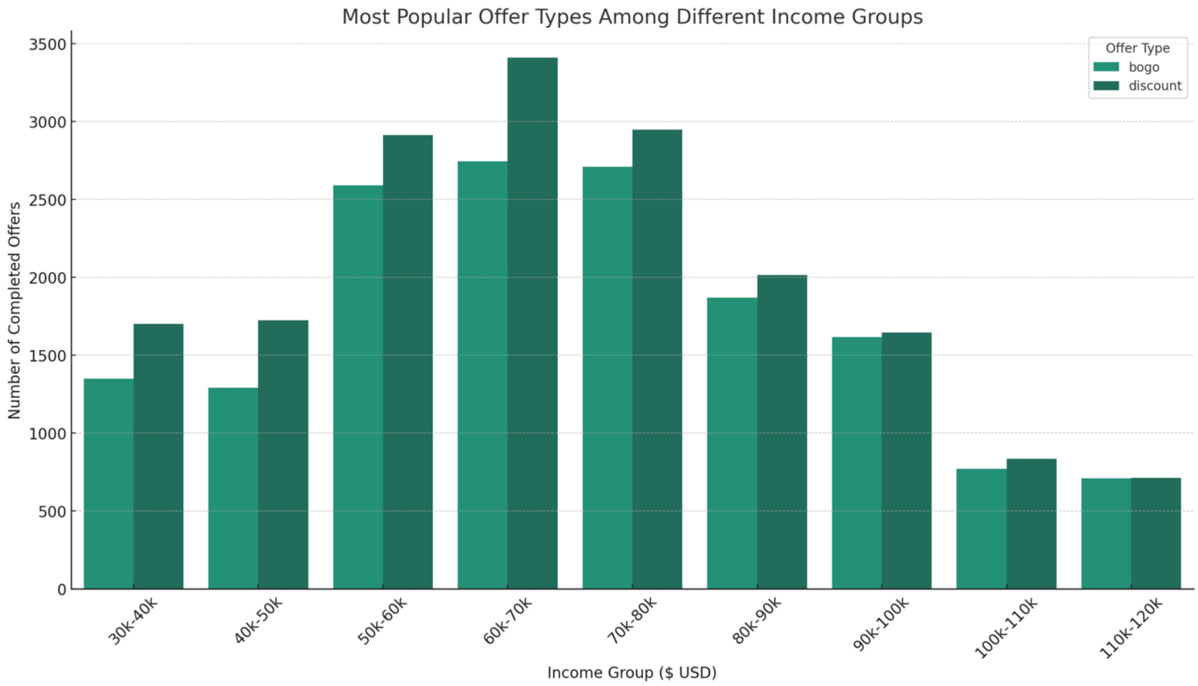
Unknown:

For customers with unknown or missing gender information, Discount offers are more popular than BOGO offers.

This suggests that Discount offers are generally more popular across all gender groups, but the preference for BOGO offers is also significant, especially among female customers and those identifying as 'Other'.

Author: Luca Zavaglia

V.6



V.6 Comments:

The bar plot illustrates the most popular offer types among different income groups, based on the number of completed offers:

For lower income groups (30k-50k):
Discount offers are more popular than BOGO offers.

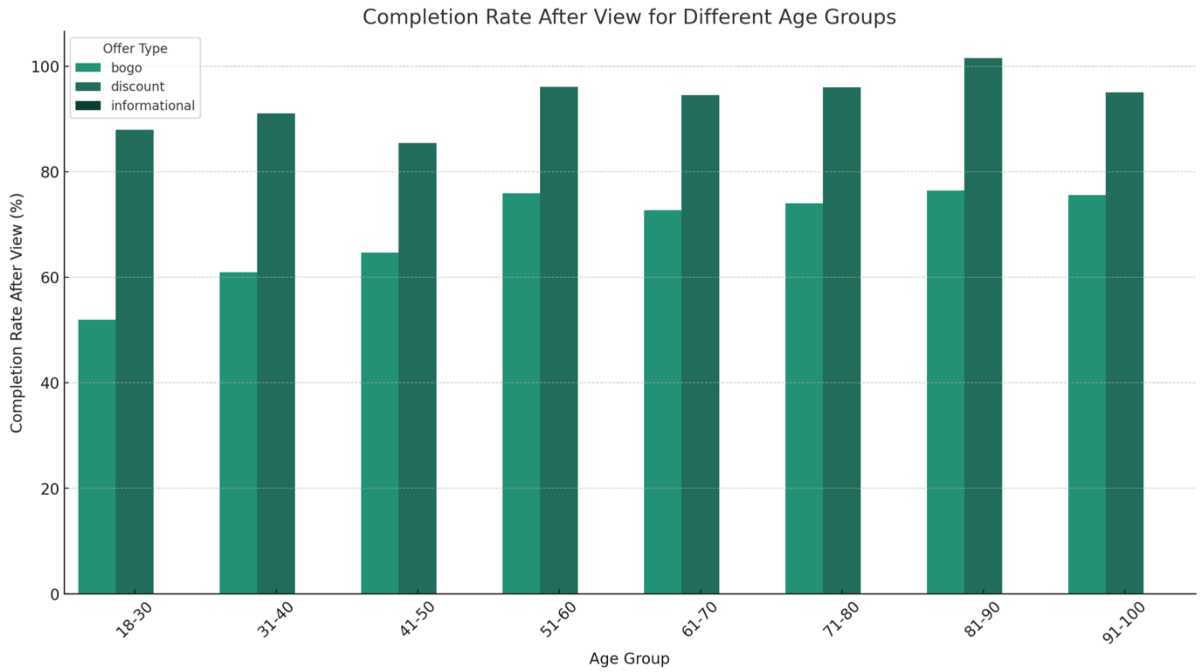
For middle income groups (50k-90k):
Discount offers are the most popular, followed closely by BOGO offers.

For higher income groups (90k-120k):
Discount offers continue to be the most popular, and the gap between Discount and BOGO offers narrows further as income increases.

This suggests that Discount offers are generally more popular across all income groups, but BOGO offers also have strong appeal, especially among middle and higher income customers.

Author: Luca Zavaglia

V.7



V.7 Comments:

The bar plot illustrates the completion rate after view for different age groups, broken down by offer type:

18-30 Age Group:
In this age group, BOGO offers have a slightly higher completion rate after being viewed compared to Discount offers.

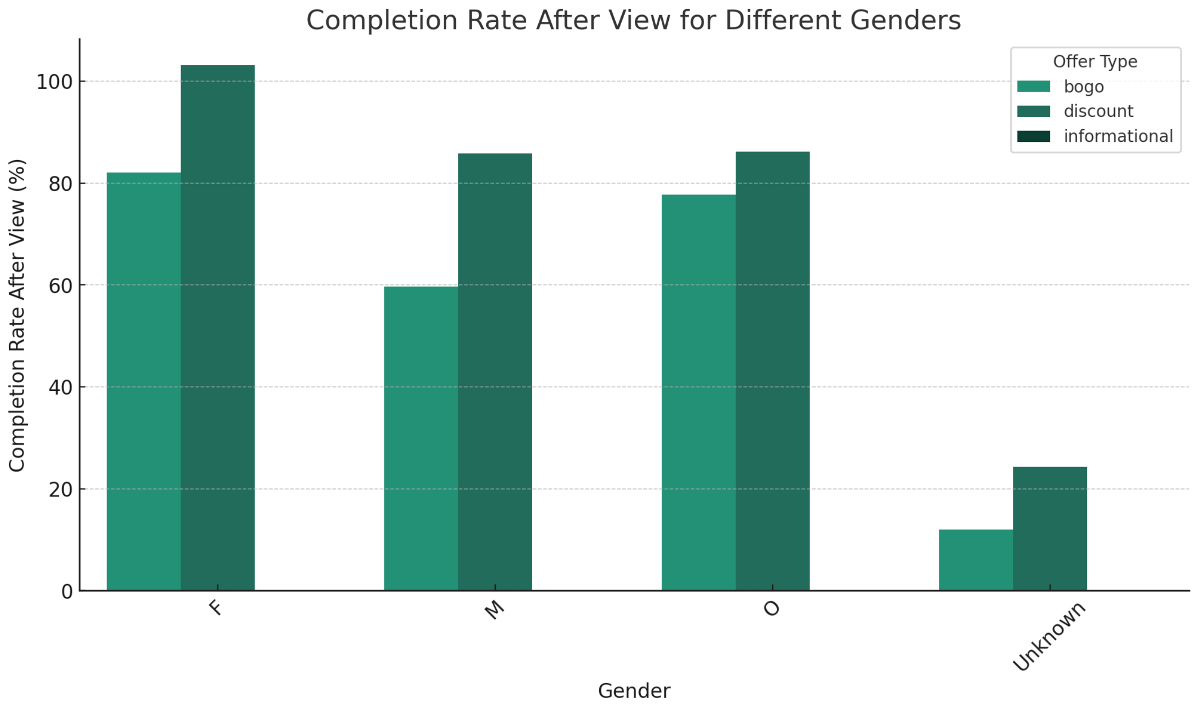
31-60 Age Groups:
For the 31-40, 41-50, and 51-60 age groups, Discount offers have a higher completion rate after being viewed compared to BOGO offers.

61-100 Age Groups:
In the older age groups (61-70, 71-80, 81-90, and 91-100), BOGO and Discount offers have similar completion rates after being viewed.

This suggests that:
Younger customers (especially 18-30) are slightly more responsive to BOGO offers after viewing them.
Middle-aged customers (31-60) are more responsive to Discount offers after viewing them.
Older customers (61 and above) respond similarly to both BOGO and Discount offers after viewing them.

Author: Luca Zavaglia

V.8



V.8 Comments:

The bar plot illustrates the completion rate after view for different genders, broken down by offer type:

Female (F):

Female customers have a higher completion rate after view for both BOGO and Discount offers compared to other gender groups. The completion rate for Discount offers is slightly higher than for BOGO offers among females.

Male (M):

Male customers have a lower completion rate after view for both BOGO and Discount offers compared to females. Similar to females, males have a slightly higher completion rate for Discount offers than for BOGO offers.

Other (O):

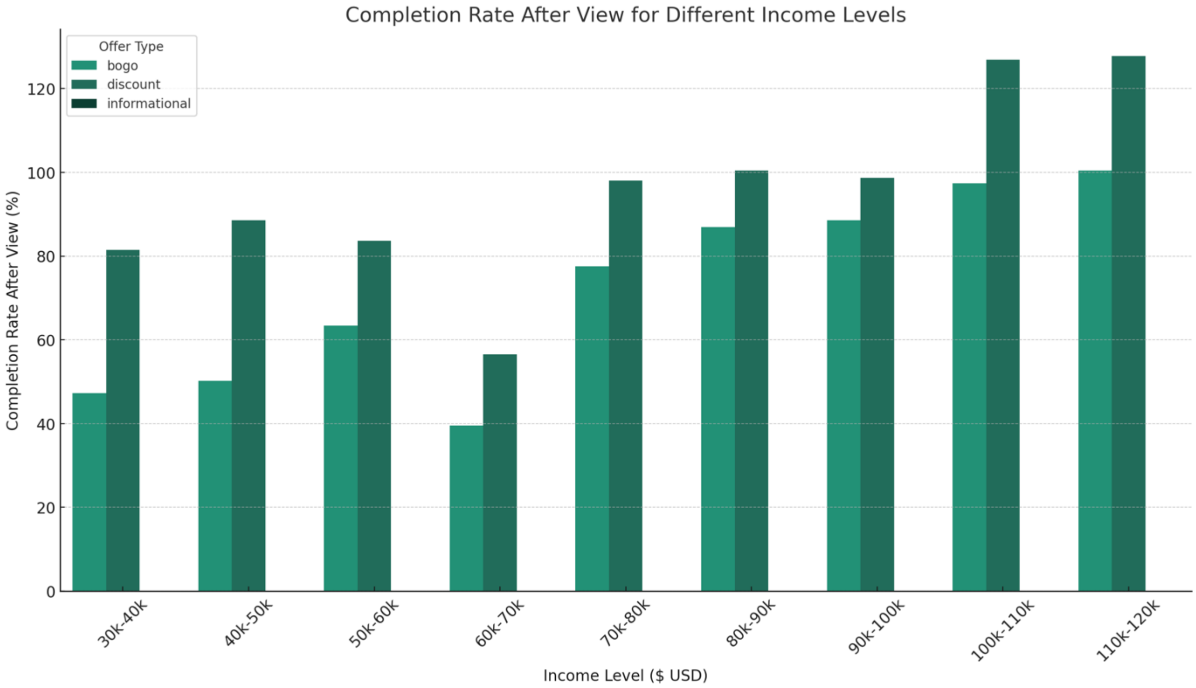
For customers who identify as a gender other than male or female, the completion rate after view is comparable to that of females, especially for Discount offers.

Unknown:

For customers with unknown or missing gender information, the completion rate after view is lower than for other gender categories.

Author: Luca Zavaglia

V.9



V.9 Comments:

The bar plot illustrates the completion rate after view for different income levels, broken down by offer type:

For lower income groups (30k-50k):
The completion rate after view for BOGO offers is similar to that for Discount offers.

For middle income groups (50k-90k):
The completion rate after view for Discount offers is slightly higher than for BOGO offers. As income increases in this range, the completion rate for both types of offers tends to increase as well.

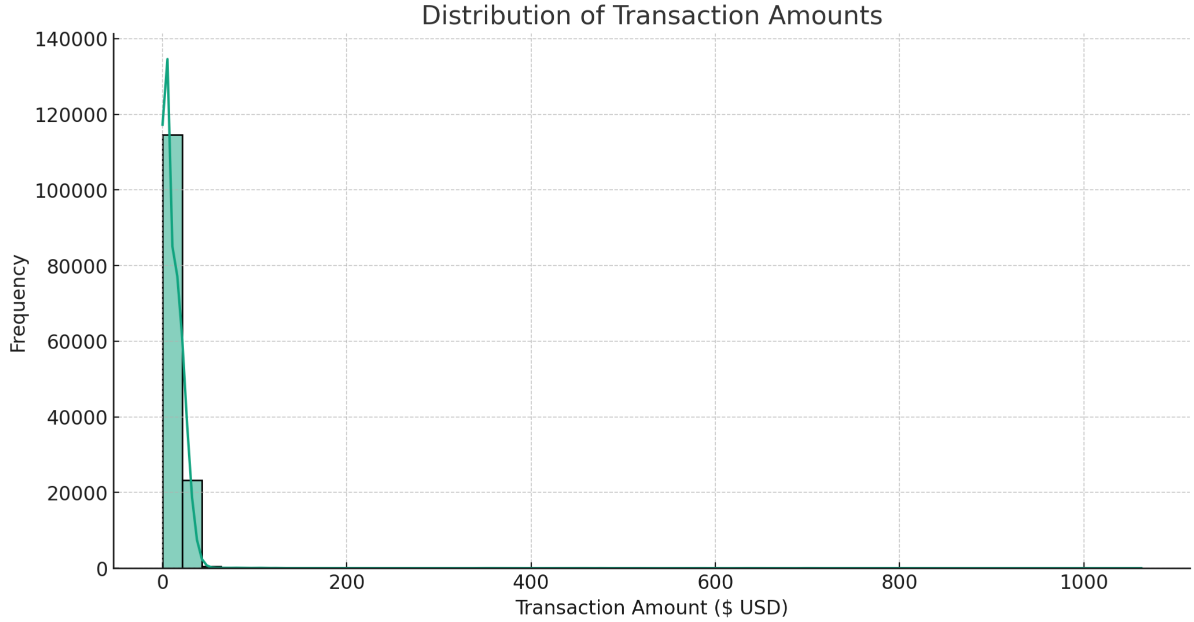
For higher income groups (90k-120k):
In these groups, the completion rate after view for BOGO and Discount offers is quite similar.

This suggests that:
For middle and higher income customers, the completion rate after viewing is high for both BOGO and Discount offers, and they respond similarly to both types of offers.
For lower income customers, they are also quite responsive to both types of offers after viewing them.

These insights are valuable for tailoring offers to different income groups. For instance, middle and higher income groups are highly responsive to both BOGO and Discount offers, while lower income groups are equally likely to respond to either type of offer.

Author: Luca Zavaglia

V.10



V.10 Comments:

The histogram illustrates the distribution of transaction amounts among customers:

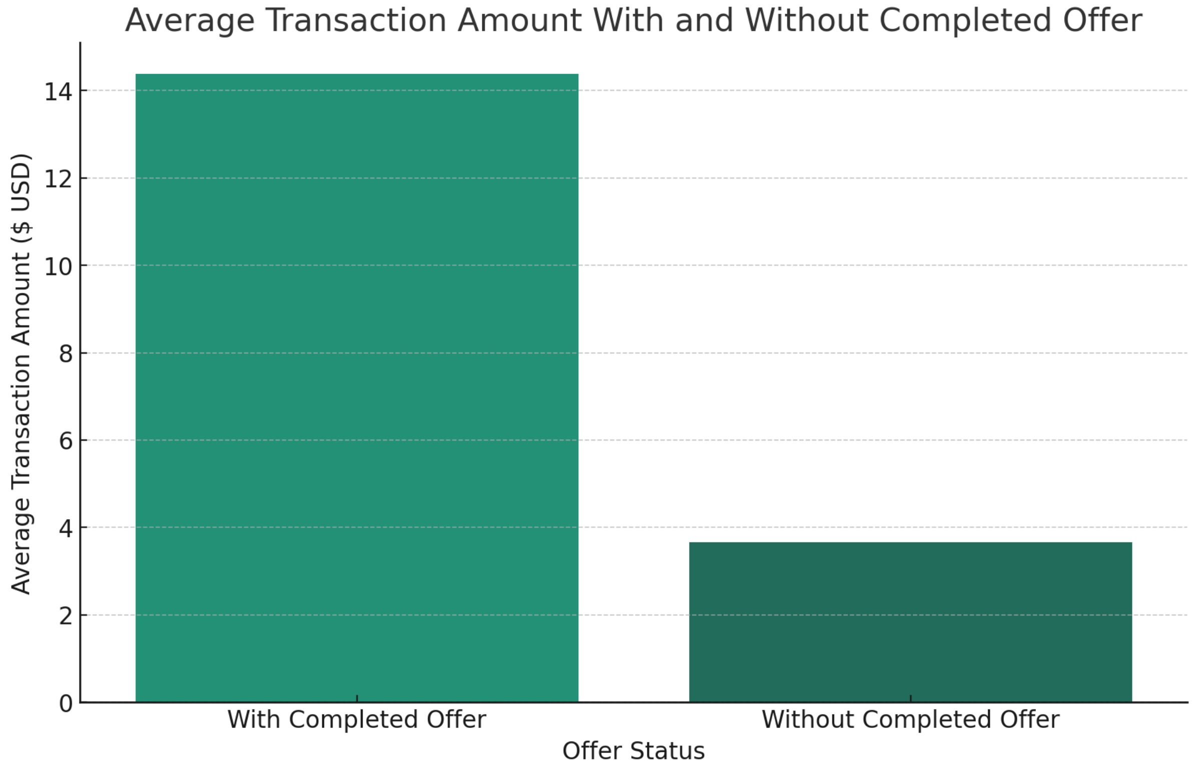
Most of the transaction amounts are clustered in the range of $0 to $30.

The distribution is right-skewed, indicating that the majority of transactions are of lower amounts, while higher transaction amounts are less frequent.

There are some transactions with higher amounts (e.g., above $40), but these are relatively rare compared to the smaller transactions.

This distribution provides an overview of the typical spending behavior of customers on the Starbucks app.

V.11



V.11 Comments:

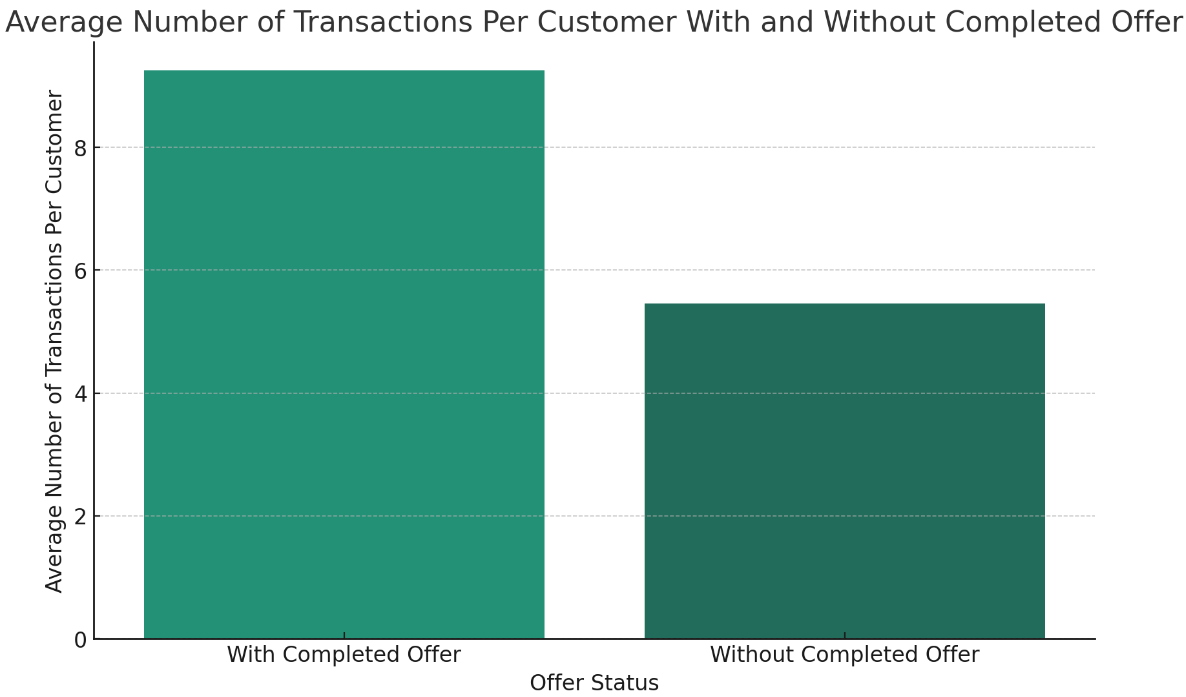
The bar plot illustrates the comparison of the average transaction amount when an offer is completed versus when no offer is completed:

The average transaction amount when an offer is completed is approximately $14.38.

The average transaction amount when no offer is completed is significantly lower, at approximately $3.67.

This suggests that customers tend to spend more in transactions associated with a completed offer compared to transactions not associated with a completed offer.

V.12



V.12 Comments:

The bar plot illustrates the comparison of the average number of transactions per customer when an offer is completed versus when no offer is completed:

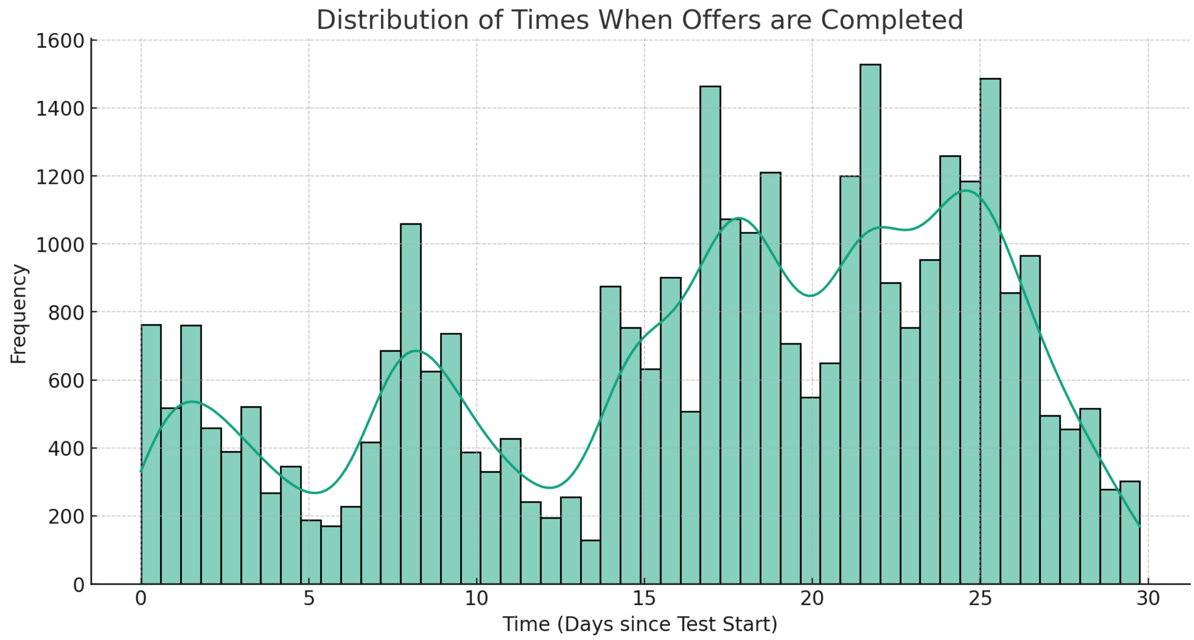
The average number of transactions per customer when an offer is completed is approximately 9.25.

The average number of transactions per customer when no offer is completed is lower, at approximately 5.46.

This suggests that customers tend to have a higher frequency of transactions when they complete an offer compared to when no offer is completed.

These insights indicate that offers, especially when completed, are associated with both a higher average transaction amount and a higher frequency of transactions per customer. This suggests that offers are effective in encouraging customers to make more frequent and higher-value purchases.

V.13



V.13 Comments:

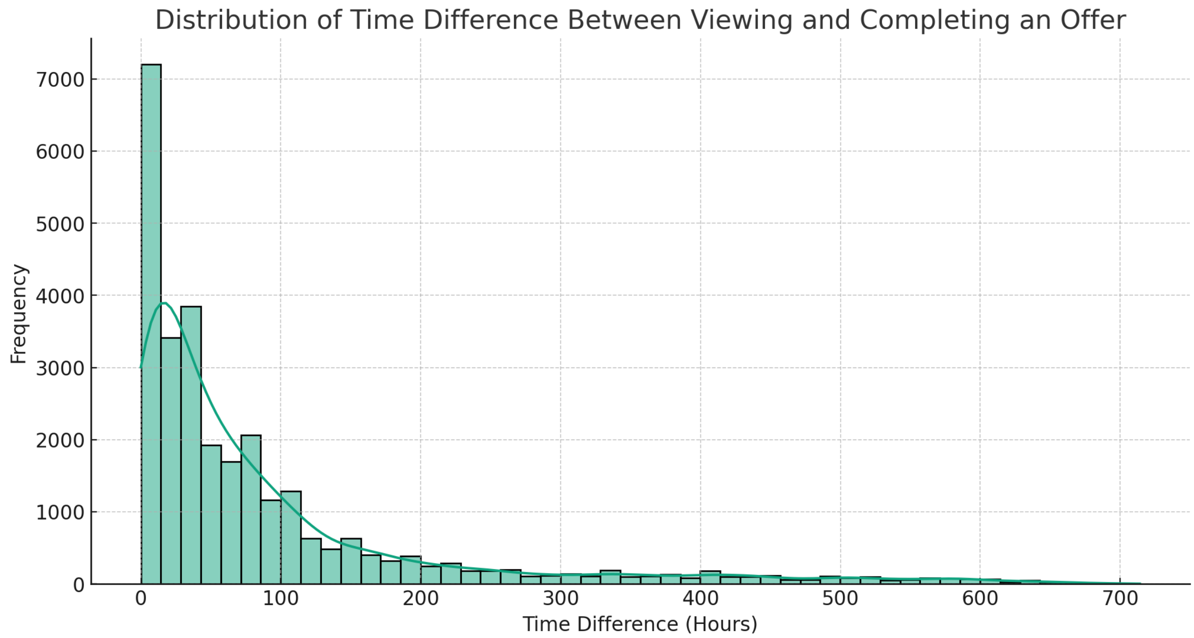
The histogram illustrates the distribution of times when offers are completed:

The distribution shows several prominent peaks. These peaks likely correspond to the times when new offers are sent out to users, prompting them to make purchases and complete the offers.

Offers are frequently completed at various times throughout the test period, with notable peaks observed around days 7, 14, 17, 21, and 24 since the test started.

This suggests that there are specific times when offers are more likely to be completed, which likely aligns with the timing of when new offers are sent out to users.

V.14



V.14 Comments:

The histogram illustrates the distribution of time differences between viewing and completing an offer:

A large number of offers are completed shortly after being viewed, as indicated by the prominent peak close to 0 hours. This suggests that many users act quickly after viewing an offer.

The distribution is right-skewed, indicating that while most users complete offers shortly after viewing them, there are also users who take a longer time to complete the offers.

There are fewer instances as the time difference increases, indicating that as more time passes after viewing an offer, users are less likely to complete it.

This analysis suggests that:
Users are generally quite responsive to offers, with a significant proportion completing the offer shortly after viewing it.

There is a wide range of responsiveness among users, with some acting almost immediately upon viewing an offer and others taking a longer time.

This information is valuable for understanding how quickly offers might lead to conversions (i.e., offer completions) and can inform the timing and frequency of sending out new offers to customers.

Author: Luca Zavaglia