Applying Iterative Design Principles to a Live Product

Name: Larni Espada
Date: December 13, 2023

KPIs & Multivariate Experiment Results

KPIs for Flyber Analyses

- For the data available, which KPI(s) best match Flyber's business model?
- How would you calculate these KPI(s) using the available event data logs?
- List other KPIs that might be important to Flyber but are not calculable based on available data

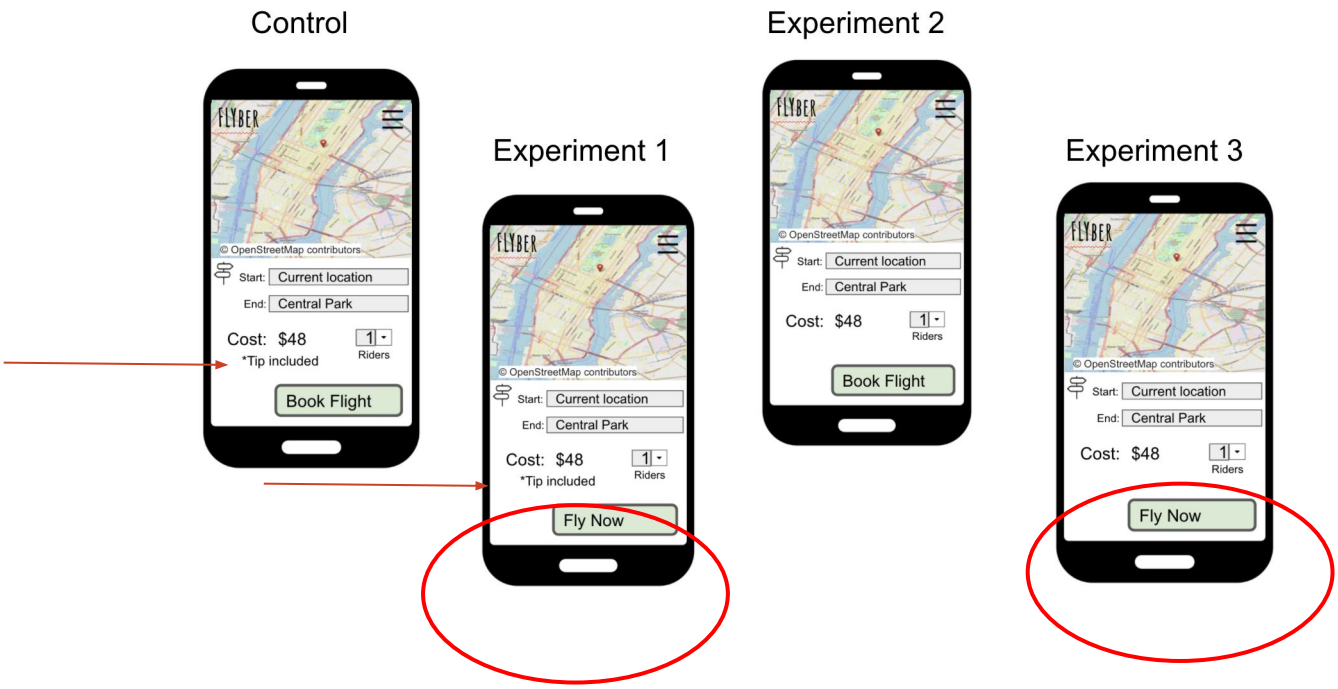
The KPIs that best match Flyber's business model based on the available data are:

- **Search Drop off** : Search is an essential metric that assesses how frequently users utilize the search feature within the Flyber app relative to the number of times they open the app. It provides insights into user engagement with the search functionality.
  Formula: Search Drop Off= (Count of "search" events) / (Count of "open" events)
- **Conversion Rate**: The Conversion Rate KPI evaluates Flyber's ability to convert users who open the app into users who initiate a ride.
  Formula: Conversion Rate = (Count of "begin_ride" events) / (Count of "open" events)
- **Daily Booked Flights**: This metric directly measures daily user engagement and conversion, which is a critical performance indicator for ride-sharing apps like Flyber.
  Formula: Daily Booked Rides = (Count of "begin_ride" events)

Other KPIs that might be important to Flyber but cannot be calculated based on the available data include:

- **First-time User Engagement**: Evaluating the engagement level of first-time users with the app by monitoring their initial actions, such as "open" and "search" events.
- **User Retention**: Measuring the percentage of users who continue to use Flyber over time to assess long-term retention.
- **Revenue Metrics**: Monitoring revenue-related KPIs such as average revenue per user, total revenue generated, and user spending patterns.
- **Customer Satisfaction**: Measuring customer satisfaction through surveys or feedback mechanisms to assess the overall user experience.

Multivariate Experiment



- Describe the elements tested during the multivariate experiment. You can use the image above when referencing the tests

The multivariate experiment involves testing three different variants of the app landing page, each with specific changes compared to the current landing page, which serves as the control group. Here are the elements tested in each variant:

Experiment 1:

- Landing Page: Replaces the "Book Flight" button with a "Fly Now" button.
- Cost Presentation: Explicitly stating that the cost includes the tip.

Experiment 2:

- Landing Page: Similar to the control group, this variant retains the "Book Flight" button.
- Cost Presentation: Specifies that the cost does not include the tip.

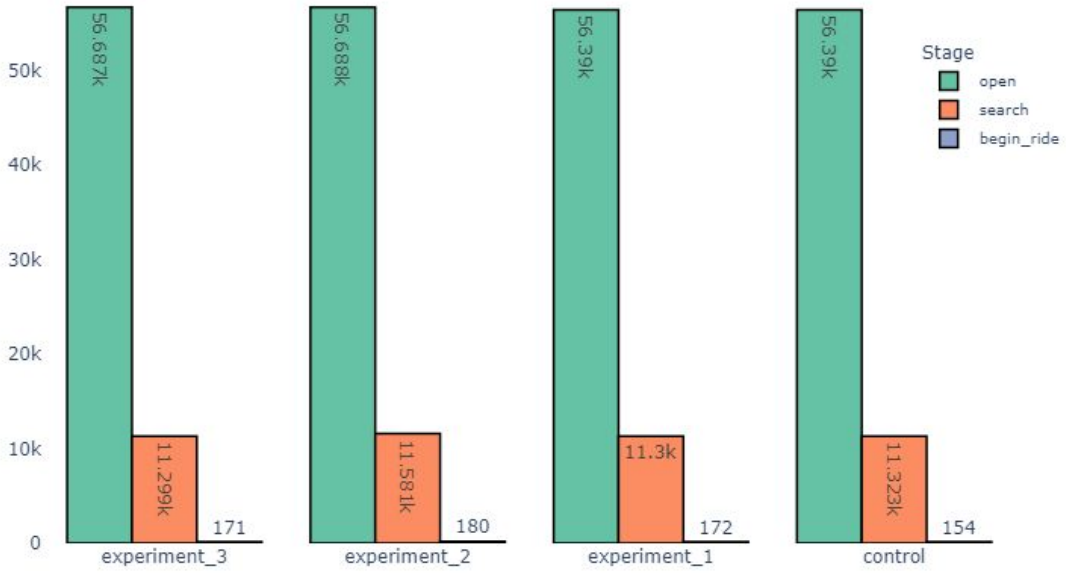
Experiment 3:

- Landing Page: Like Experiment 1, this variant includes the "Fly Now" button.
- Cost Presentation: Similar to Experiment 2, specifies that the tip isn't included.

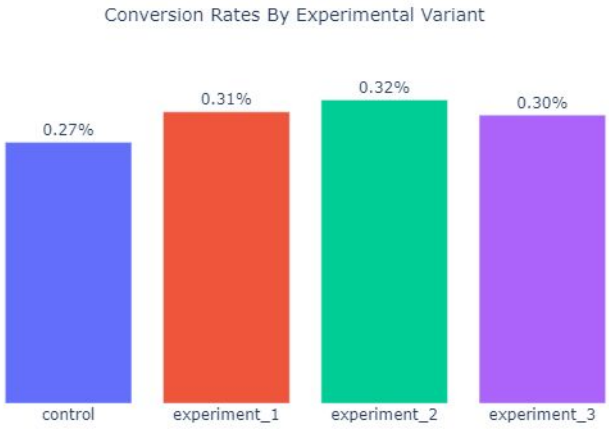
These experiments are designed to evaluate the effects of the proposed alterations in the landing page and cost presentation on user behavior and engagement in comparison to the Control group, which features a "Book Flight" button with tip included in cost. It's worth noting that all the variants, including the Control Group, collectively constitute a 2x2 Factorial Design.

Multivariate Test Results:

- Provide a visual representation of the impact of the experiment on the conversion rate of users booking a flight (out of all users opening the app)



The conversion rates (% booked flight out of users who opened the app) for the different experimental variants hover around 0.3% to 0.32%, which appears to be just slightly higher than the control group's rate of 0.27%.



Multivariate Test Results:

- Determine if there was a significant difference between the experiments and control states.

The z-test for two proportions is a statistical test that assesses whether there is a significant difference between two proportions from independent groups. It calculates a z-score to measure the deviation between the sample proportions and estimates the population proportions. *This z-test is appropriate when the sample sizes are sufficiently large (n > 30).*

Test Statistic for Two Proportions (with H_0: $p_1 = p_2$)

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\bar{p}\bar{q}}{n_1} + \frac{\bar{p}\bar{q}}{n_2}}}$$

where $p_1 - p_2 = 0$ (assumed in the null hypothesis)

where $\bar{p} = \frac{x_1 + x_2}{n_1 + n_2}$ (*pooled* sample proportion) and $\bar{q} = 1 - \bar{p}$

In the context of the experiment, the z-test results below help us understand the significance of the differences in conversion rates between the control group and the experimental variants. The "NS" (Not Significant) label indicates that the observed differences in conversion rates are not statistically significant (p-value < 0.016, Bonferroni adjusted p-value threshold).

Overall, these results suggest that the variations introduced in the experimental groups do not have a substantial impact on conversion rates compared to the control group.

Control vs experiment_1 two sided z-test: z = 0.9984, p value = 0.3181, NS
Control vs experiment_2 two sided z-test: z = 1.3765, p value = 0.1687, NS
Control vs experiment_3 two sided z-test: z = 0.8969, p value = 0.3698, NS

Since all three experiments did not outperform the Control group, there is no compelling evidence to support the expansion of any specific experiment at this stage. However, it's crucial to consider additional factors and conduct further analysis before arriving at a final decision. Expanding experiments should involve a thorough assessment of both statistical significance and practical significance, along with a deeper exploration of user behavior and preferences.

Consider also refining and redesigning the experiments based on current results. A well-planned, improved test can provide more valuable insights and potential strategies for enhancement.

Funnel &
Cohort Analyses

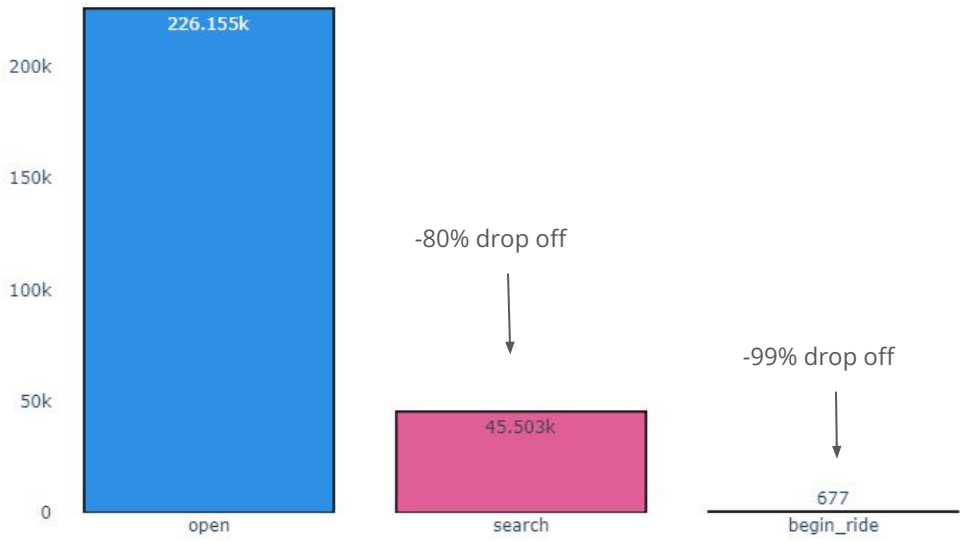
User Funnel

Identifying the different stages the user funnel

- Based on the event types in the data provided, list the 3 or more steps a user can take from opening the app to final booking of a ride
- Provide a graph showing the funnel from step to step, including drop off rates.

Based on the event types in the data, a user can take the following steps from opening the app to final booking of a flight:

Open the app: The initial step where users open the Flyber app.
Search for a destination: Users can search for their desired ride destination within the app.
Begin the ride: The final step where users initiate and book a flight..

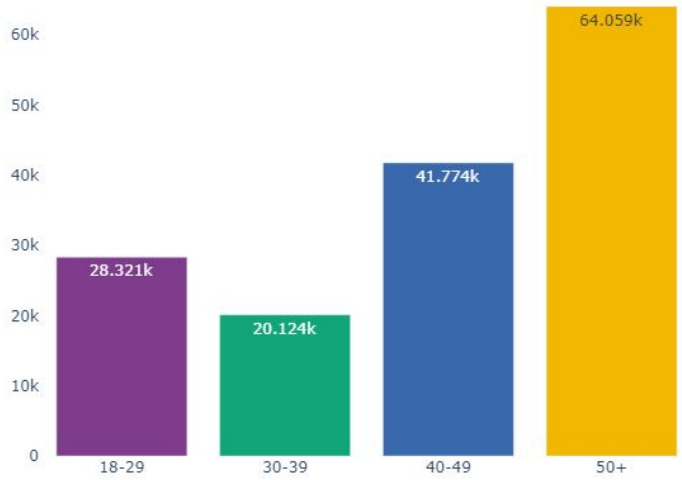


It's clear from the observed drop-off rates that there are significant challenges affecting user engagement and conversion within the app. The data reveals a substantial 80% drop-off from the initial app launch to the point of searching for destinations, followed by a 99% drop-off from the search stage to completing a flight booking. These figures suggest underlying issues in the user journey and raise concerns about the app's overall conversion funnel.

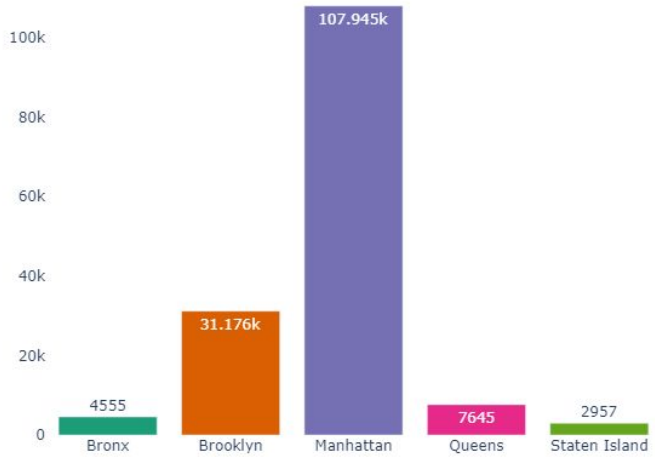
User Segments

- Identify 2 demographic attributes present in the data that allow for segment analysis
- For each demographic attribute, provide the number of users in each segment group
- For each demographic attribute, identify the segment group with the largest number of users

- ***Age group 50+ represents largest segment at 42% (64K).***



- ***Manhattan borough represents largest segment at 70% (108K).***



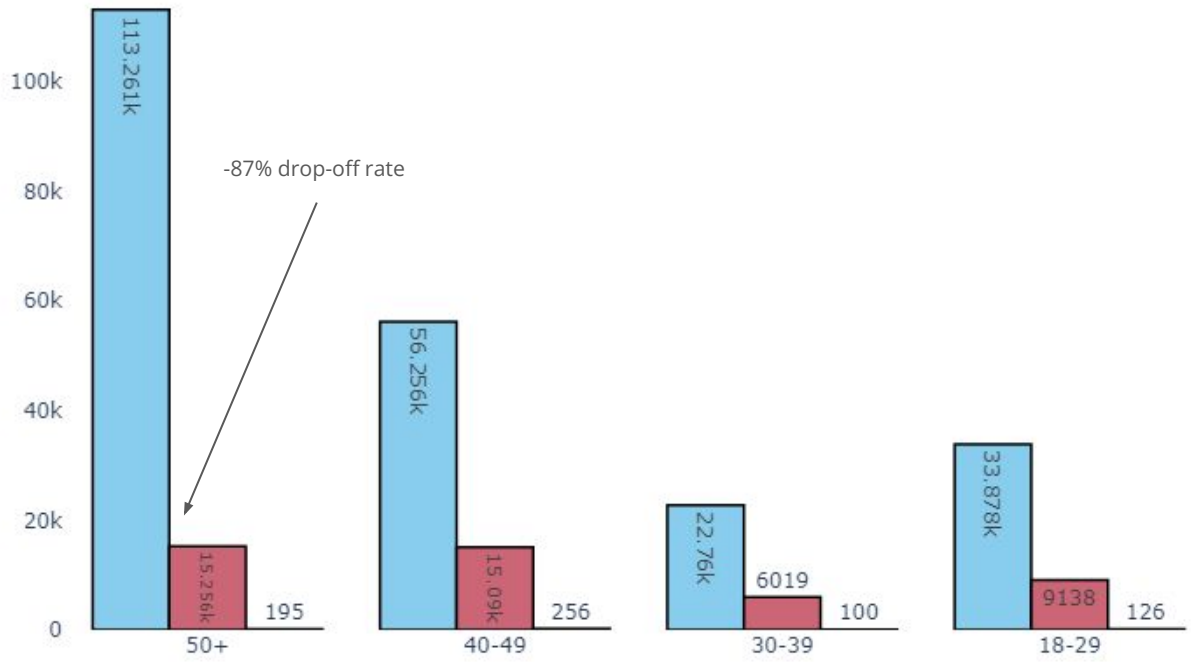
Segment Analysis of Funnel

Identify Opportunities for Improvement

- Perform a funnel analysis by segment for all identified demographic attributes and describe the results
- If underperformance for a segment in an attribute is identified, add a visual showing the average funnel conversion by segment group for that demographic

➔ In the demographic funnel analysis, the **50+ age group** exhibited the most significant drop-off, with an **87% decrease from opening the app to searching**. In contrast, all other age groups experienced an 73% drop-off. The drop-off rates from searching to booking a flight were consistent across age groups.

	18-29	30-39	40-49	50+
Open App	33,878	22,760	56,256	113,261
Drop-off Search	-73.0%	-73.6%	-73.2%	-86.5%
Drop-off Booking	-98.6%	-98.3%	-98.3%	-98.7%
# Conversions	126	100	256	195



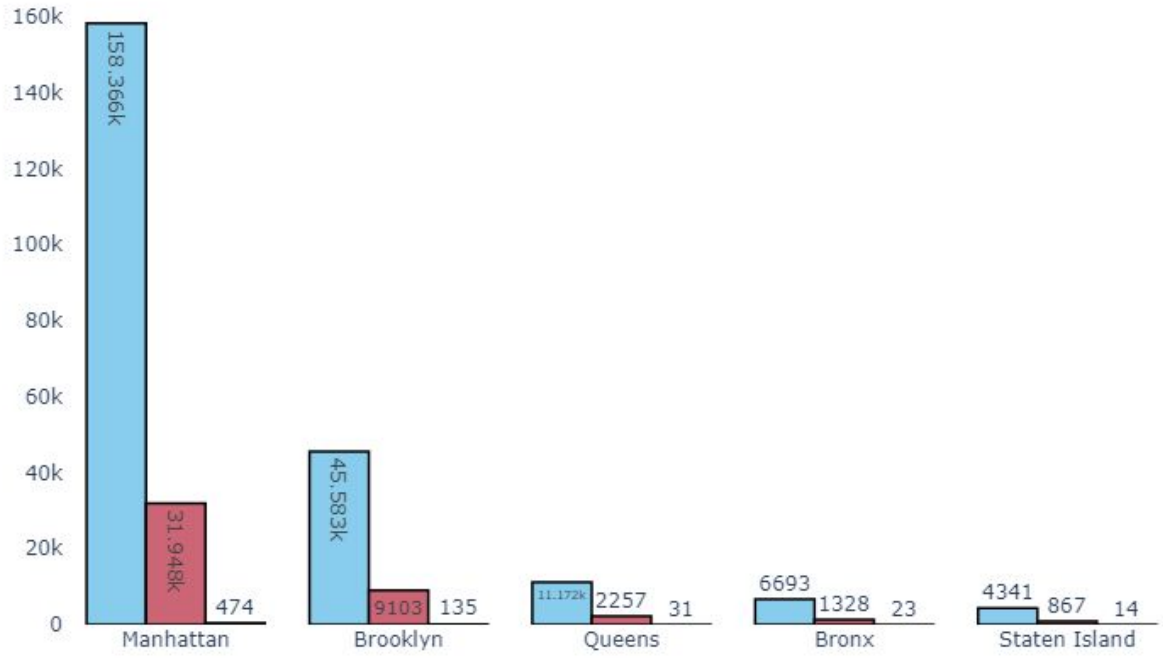
Segment Analysis of Funnel

Identify Opportunities for Improvement

- Perform a funnel analysis by segment for all identified demographic attributes and describe the results
- If underperformance for a segment in an attribute is identified, add a visual showing the average funnel conversion by segment group for that demographic

➔ Regarding borough segments, the drop-off rates from opening the app to searching (-80%) and from searching to booking a flight (-98%)were similar across all boroughs.

	Bronx	Brooklyn	Manhattan	Queens	Staten Island
Open App	6,693	45,583	158,366	11,172	4,341
Drop-off Search	-80.2%	-80.0%	-79.8%	-79.8%	-80.0%
Drop-off Booking	-98.3%	-98.5%	-98.5%	-98.6%	-98.4%
# Conversions	23	135	474	31	14



Hypothesis & Next Steps

Review Qualitative Data

- Read user interviews to understand “why” any funnel under-performance seen in Step 2 might occur
- List your hypothesis for what customer need is being under-served
- Provide 3 or more quotes as evidence for this hypothesis

Based on the drop-off rates observed in Step 2 of the demographic funnel analysis, it's important to delve into the user interviews to understand why these underperformance occur. Here's a hypothesis for the customer need that might be underserved:

Hypothesis: The 50+ age group may be experiencing difficulties in using the app due to a lack of user-friendliness or specific features catering to their needs. This age group might require a more intuitive and simplified app interface to continue using it effectively.

Supporting Quotes from 50+ age group:

- "I just hail a taxi or tell my phone to call a cab to go to a certain address (I'm always on the phone, so I just use voice commands with my phone most of the time)" - **Challenge: Usability and Habitual Use of Alternatives**
- "Drive my car or call a taxi service." - **Challenge: Habitual Use of Alternatives**
- "I have a personal car service on call. My assistant books Flyber whenever I'd be traveling during peak NYC traffic hours. Time is money and Flyber saves me time! But I let my assistant actually book the Flyber because the first few times I tried booking, the instructions were too small." - **Challenge: Usability and Habitual Use of Alternatives**
- "Honestly, I thought about using Flyber to surprise my grandson or granddaughter with a visit to one of their sporting games. Luckily my daughter was around to help me book the ride. I usually just use Uber because it remembers my addresses and has all my favorite places saved, so I guess I always just open that up since it is so convenient and saves me time. Though now that I say that, I really should use Flyber again since it would save more time when it comes to fighting traffic!" - **Challenge: Usability and Convenience**
- "I call up our local pilot, Bob. He's not always available but I don't need to fiddle around with an app and hitting tiny buttons. He knows where I tend to be and where I want to go." - **Challenge: Usability and Transportation Alternatives**

These quotes suggest that interviewed users, especially in the 50+ age group, prioritize convenience, ease of use, and familiarity. The current app interface might not be meeting their specific needs or could have usability issues, which could explain the higher drop-off rate observed in this demographic.

*7 out of 12 interviewed are in 50+ age group

Suggested Features & Experimentation Plan

- Share your hypothesis using the following format: We believe [observed quantitative effect] Because [hypothesized user “why”] And that by [general change/opportunity for Flyber to improve] for [targeted cohort] we will see [expected effect]

Hypothesis:

We believe that the high drop-off rates observed in the 50+ age group, particularly during the transition from opening the app to searching for destinations, are due to usability challenges and a lack of user-friendly features that cater to their needs. Older users may find the current app interface less intuitive and may prefer more familiar and convenient options.

And that by enhancing the app's user-friendliness, providing features like voice commands, guided onboarding, personalized assistance, and other intuitive features targeting these improvements specifically to the 50+ age group, we will see a significant reduction in drop-off rates within this demographic, resulting in improved user engagement and retention among older users.

Expected Effect:

We expect that by addressing the usability challenges faced by the 50+ age group and providing a more tailored and user-friendly experience, we will observe a decrease in the drop-off rate from opening the app to searching for destinations within this demographic. This should lead to a higher level of user engagement and retention among older users, ultimately contributing to Flyber's success in serving a more diverse user base.

Suggested Features & Experimentation Plan

- Suggest 2 or more features that would match your hypothesis and determine a plan for multivariate testing, including describing the control and experimental conditions
- Determine who should be exposed to the experimental changes

Features for Multivariate Testing:

For this multivariate testing, we have chosen two key features: App User Interface and Save Prior Addresses and Destinations. These features have been selected due to their potential for high impact and wide-reaching benefits, while also being relatively easy and cost-effective to prototype (RICE framework).

- ➔ **Control Group**: This group will not be exposed to either of the two new app features, serving as a baseline for comparison.
- ➔ **Variant 1**: This group will be exposed to the new app user interface only, allowing us to assess the impact of this UI change on user behavior and engagement.
- ➔ **Variant 2**: Users in this group will have the ability to save prior addresses but will still be using the current app user interface. This variant will help us evaluate the impact of the Save Prior Addresses and Destinations feature independently.
- ➔ **Variant 3**: This group will be exposed to both new app features: the updated user interface and the ability to save prior addresses. This will enable us to understand the combined effects of these changes on user interaction and conversion rates.

Plan for Multivariate Testing:

- Users in the 50+ age group will be randomly selected and divided into control and three experimental groups listed above. Group assignments will be predetermined.
- An email campaign will be initiated to direct users in the experimental groups to visit the app and explore the new features.
- The testing period will span one month to collect sufficient data for analysis. . **Data tracking will continue for an additional month post-testing to monitor any sustained effects and gather insights into long-term user behavior.**

This approach ensures a controlled and structured testing environment for evaluating the impact of these new features on user engagement and usability within the 50+ age group. **The predetermined group assignments will minimize online handling and maintain the integrity of the testing process.**

Suggested Features & Experimentation Plan

- List any additional metrics that would be helpful to collect from your suggested features

In-Test Phase Metrics:
During the testing phase, we will collect a range of metrics to assess the impact of the implemented changes:

- **Drop-off Rates**: We will diligently track drop-off rates from the moment users open the app to when they initiate searches for destinations. These rates will be meticulously recorded for both the control and experimental groups within the 50+ age category. This metric will help us identify any potential issues in the user journey.
- **Successful Conversions**: To gauge the effectiveness of the alterations, we will monitor the conversion rate at the end, specifically focusing on completed bookings, among users aged 50+. This metric will provide insights into the functionality and efficiency of the implemented changes and their ability to drive desired user actions.

Additional Metrics:
In addition to the core metrics mentioned above, we will also consider the following supplementary metrics:

- **Engagement and Retention Rates**: User engagement and retention rates within the 50+ age group will be closely monitored during the post-testing period. This longitudinal analysis will enable us to assess the sustained impact of the implemented features on user behavior, providing insights into their long-term effectiveness.
- **ROI Analysis**: To comprehensively evaluate the campaign's financial aspects, we will conduct a detailed ROI analysis. This analysis will encompass a thorough assessment of the costs associated with implementing the new features, revenue generation directly attributed to the changes, cost reduction in relevant areas, and net profitability. The resulting ROI figure will offer a clear indication of the campaign's financial success and sustainability.
- **User Feedback**: Additional qualitative insights will be gathered through in-app surveys designed to understand the perceived usability and user-friendliness of the introduced changes. User feedback is invaluable in uncovering user sentiments and preferences, allowing us to fine-tune the features based on real user input.

By including these metrics and providing a more detailed explanation, your assessment will offer a comprehensive view of the evaluation process, encompassing both quantitative and qualitative dimensions, and financial implications.