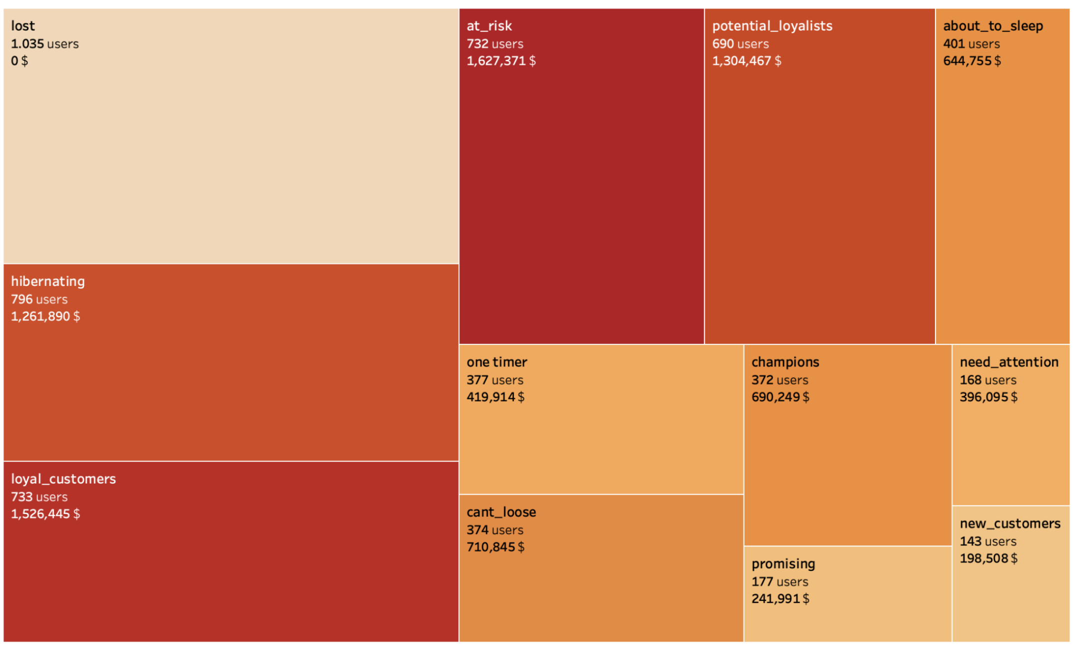
Report: TravelTide Segmentation

Introduction: The primary challenge for TravelTide is to segment its users effectively. This segmentation will enable better understanding and targeting of customers based on their behaviours.
Considering the relatively limited data available, we have thought of a methodology based on few metrics: from their most general to most personal connotation.
That way, we are able to introduce a semi-randomness to balance with the biassed inferences we would inject while defining a group according to few observations.
The proposed logic would also allow for scalability, as the more users would commit to TravelTide services, the more we would be able to define nuanced groups and adapted perks.
Additionally, we also performed an RFM (Recency, Frequency, Monetary) classification that provides complementary insights into the overall health of TravelTide users.

Data and EDA: To note, that our segmentation includes users who have consulted TravelTide services at least 7 times since April 1, 2024. This criterion results in around 6,000 users for our segmentation analysis.
We also had to compose with data inconsistencies that we adjusted in the following manner:

- Negative counts of nights: imputed with 0
- 0 nights for completed trips: imputing the difference between departure and return flights
- Removing the duplicate trip references for cancel trips at sessions level

RFM Classification: Users are first categorised based on their activity on our platform, following scores from Recency, Frequency, Monetary metrics.
By observing how often they use our service; the last time they used it; and their economic contribution, we are able to segment them into groups with associated perks.

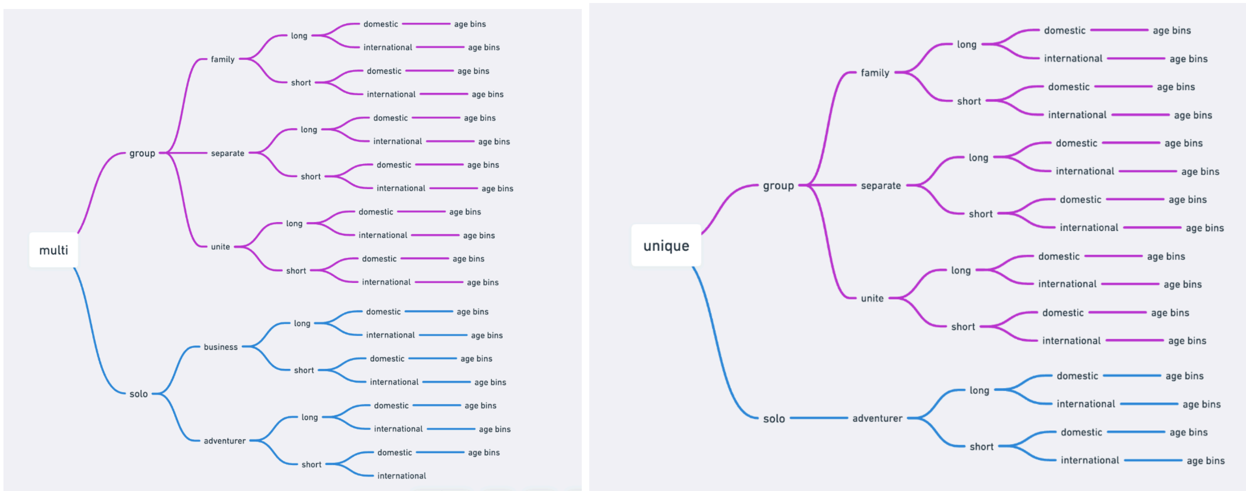


- Users distribution within the RFM segmentation

From this segmentation, one can observe that certain groups should be addressed in priority, as the users, despite representing a large economical benefit, are close to being considered inactive.

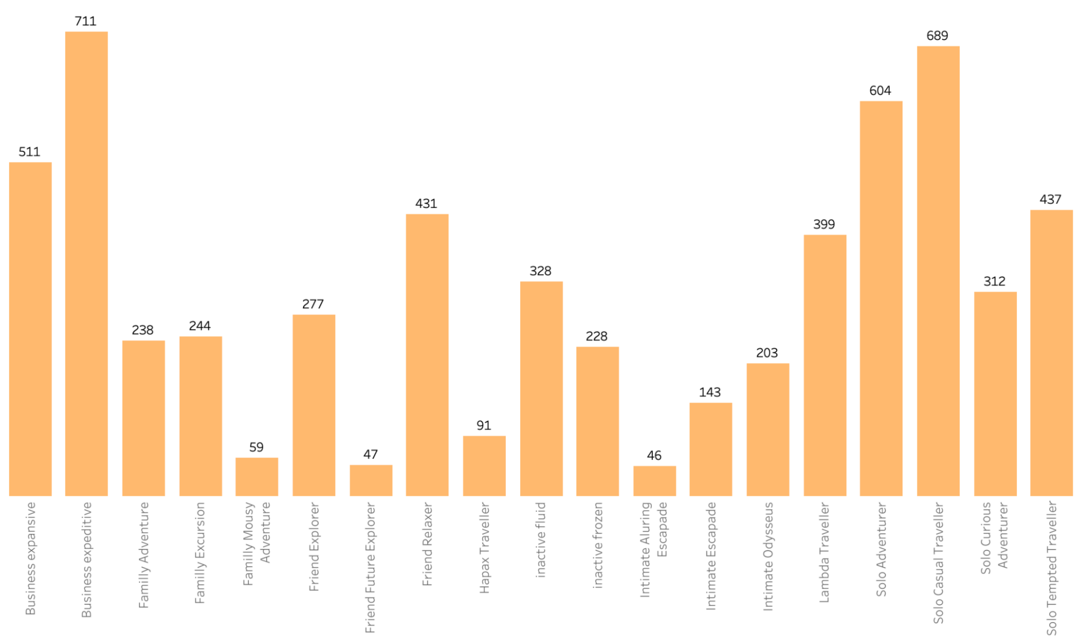
Behavioural Segmentation: Our main segmentation defines groups based on inferred behaviours from a limited set of metrics. The division in groups happens through a branching system that divides users with the following logic:

- Whether a user has booked one, multiple, or no trips.
- Solo or group travellers (determined by the number of reserved seats).
- Type of travellers (based on room usage, trip duration, and frequency).
- Long or short trips (determined by the number of nights spent on-site).
- Age groups, divided in three buckets.



- Branching for segmenting users with multiple or unique trips

These groups could be further distributed, additionally considering segments such as: domestic/international travels or seasonality.
Since our user base isn't so large, we decided on establishing a threshold in this segmentation.
That way, we can compose with groups that are large enough to be considered as a sensible population.
Thus, we only considered group divisions until the 'long' or 'short' trips segment.
19 groups were created via this methodology and as many perks were associated with each of their labels.



- Users behaviour distribution

Additional Insights: Since we are aware that our method does not consider metrics that could nonetheless be of importance in inferring a behaviour to a user and categorising them, we are providing two additional metrics, at user level, to consider at one's discretion.
Thus, each user is associated with a ratio for their discount usage: 'Cold Opportunist', 'lukewarm Opportunist' or 'Hot Opportunists' labels, corresponding to users who use no or almost no discount; users who use at least 1 discount per trip; users who uses 2 or more discounts per trip.
This way, one may consider the possible effectiveness of a discount offer on a particular user.
Similarly, we added information about the season and month at which users are most active.
This could help in deciding when to reach a peculiar user and maximise chance of success in offers acceptance and trip booking.

Conclusion: While we trust the proposed method to offer solid behavioural inferences, it still partly relies on interpretation, hence is subject to bias.
Thus, the next step should include testing, to confront our findings with real, collected data on the reaction of our users to the proposed perks. This could take the following shapes:

- Conducting A/B tests to compare proposed perks against randomly assigned perks.
- Entice users for polling to gather feedback on travel habits and preferred perks.

On a promising note, some general observations came up during EDA that could inspire TravelTide if they are seeking to expand and diversify their user base.
If it is a known socio-cultural behaviour that women travel more than men, the distribution of these populations in TravelTide users base still falls below standard average. We would either encourage to focus our market on women or to develop research and to understand this men deficit and how to interest them in travelling.
A less drastic difference can be observed between international and domestic trips, as ⅔ of trips are set domestically. If that proves of interest, one could investigate on why this ratio and potential to develop international offers.
We also believe that the scaling of our method would allow a more thorough understanding of our users regarding seasonality(with fewer trips made in Autumn and Winter), discount usage, trips that only comprise booking of hotel or flight; preferences such as favourite airlines, destinations or hotels.
This could help us develop more nuanced persona and perks that will answer more precisely our users' needs and demands.

* - 80% of travel decision are taken by women, though, at TravelTide they represent almost 90% of the user base - https://roarmedia.com/blog/empower-women-travel-capturing-the-female-travel-market-in-2024/#:~:text=According%20to%20a%20study%20by,booking%20accommodations%20and%20planning%20activities.