Defining and Understanding Passive LTV in an Online Dating Network

Who to Target?

What factors impact passive LTV and how can these findings impact Bumble’s future work?

How should Bumble define an accurate measure of passive LTV?

Understanding Passive LTV

Defining Passive LTV

Step 1

Project Sponsors: Laura Mitchell, Luca Cerone, Nicola Gno, William Fletcher

Faculty Advisor: Dr. Amy Farahat

MBa Students: Amy Ho & Siqi Wu

Bumble operates on a freemium model and generates revenue through its premium subscription service and paid premium features. Across Bumble’s active users, approximately 1.8 million are payers and the total value of these payments can be attributed to ‘passive users’ – users who are shown as options to swipe on during an app session.

Therefore, Bumble needs to understand the value that passive users bring to the platform. Our problem statement can be split into 2 parts:

- How should Bumble define an accurate measure of passive LTV?
- What factors impact passive LTV and how can these findings impact Bumble’s future work?

Results

We established a company-wide definition of Passive LTV for Bumble users, agreed by both the data science team and stakeholders.

We provided the product revenue team with the ability to look at average PLTV over time and consumables

We also confirmed that User Demographics are extremely important to a user being valuable to Bumble’s platform for use in future marketing campaigns.

We provided Bumble with the following deliverables:

- Passive LTV Documentation
- Full research process on Passive LTV definition
- Modularized Code/Models
- 3 Python Scripts with a README file for reproducibility
- Experimentation Recommendations
- Potential A/B testing experiments for the future

Next Steps

Following our project, Bumble will continue to validate our Passive LTV definition and test if our findings are generalizable. This includes:

- Introducing more premium interactions into passive LTV definition
- Performing A/B testing on different users to track their responses to product pricing strategies
- Productionizing our pipeline with appropriate data ingestion

Timeline

Admin

Data Exploration

Data Access

EDA

Initial Definition

Baseline Clustering

Definition Iteration

Factors & Stakeholder Agreement

Deliverables

Methodology

Step 1

We proposed the idea that passive value can be defined as the proportion of an active user’s investment on the passive user.

Step 2

After iterating through 4 candidate definitions, we chose the one that included regular Yes Votes, SuperSwipes and Beeline matches. The steps for calculating passive LTV are:

1. Identifying features as different users value features differently

2. Allocating based on influential ties between users

3. Aggregating Passive LTV by adding up total passive value received

Bumble operates on a freemium model and generates revenue through its premium subscription packages and consumables.

Criteria for Sample Scope:

- Registered between 2022.1.1 and 2021.4.30
- Profile location in Los Angeles
- Non spambot/test users

Final Sample Size: 32.4k users

Defining Passive LTV

Data & EDA

The data we used can be divided into 3 parts:

- User Data
  - Includes demographics, lifestyle habits and personal interests
- Activities Data
  - Daily aggregation of user activities, event log of voting results
- Revenue Data
  - Transactions on Bumble subscription packages and consumables

High influence & likely to pay

High influence & unlikely to pay

Limited Opportunity

Low influence & unlikely to pay

Low influence & likely to pay

Rising star

Low

High

Internal Popularity

External Popularity

Metrics

Score

Precision

AUC

0.70

0.58

0.79

Model Explainability

Shapley Values

Week since registration

Average Passive LTV over Time

User Data

Activities Data

Revenue Data

Features:

- User demographics
- Early activity indicators
- Current passive LTV

Results:

- Across a user’s lifetime, out-of-sample model performance is consistent

Model 1. Predicting future Passive LTV with early activity indicators

Features:

- User demographics
- Early activity indicators
- Current passive LTV

Input

Target

Results:

- Out-of-sample R-squared in predicting passive LTV

Model 2. Identifying key drivers of the most valuable user group

Features:

- Multiclass labels
- Best User Group
- High influence & likely to pay
- High influence & unlikely to pay
- Loyal Member
- Low influence & likely to pay
- Limited Opportunity
- Low influence & unlikely to pay

Results:

- Out-of-sample R-squared in predicting passive LTV

We also confirmed that Age, Height, Gender, Exercise, Company are extremely important to a user being valuable to Bumble’s platform for use in future marketing campaigns.

We will continue to validate our Passive LTV definition and test if our findings are generalizable. This includes:

- Introducing more premium interactions into passive LTV definition
- Performing A/B testing on different users to track their responses to product pricing strategies
- Productionizing our pipeline with appropriate data ingestion

We established a company-wide definition of Passive LTV for Bumble users, agreed by both the data science team and stakeholders.

We provided the product revenue team with the ability to look at average PLTV over time and consumables

We also confirmed that User Demographics are extremely important to a user being valuable to Bumble’s platform for use in future marketing campaigns.

We provided Bumble with the following deliverables:

- Passive LTV Documentation
- Full research process on Passive LTV definition
- Modularized Code/Models
- 3 Python Scripts with a README file for reproducibility
- Experimentation Recommendations
- Potential A/B testing experiments for the future

Next Steps

Following our project, Bumble will continue to validate our Passive LTV definition and test if our findings are generalizable. This includes:

- Introducing more premium interactions into passive LTV definition
- Performing A/B testing on different users to track their responses to product pricing strategies
- Productionizing our pipeline with appropriate data ingestion

We established a company-wide definition of Passive LTV for Bumble users, agreed by both the data science team and stakeholders.

We provided the product revenue team with the ability to look at average PLTV over time and consumables

We also confirmed that User Demographics are extremely important to a user being valuable to Bumble’s platform for use in future marketing campaigns.

We provided Bumble with the following deliverables:

- Passive LTV Documentation
- Full research process on Passive LTV definition
- Modularized Code/Models
- 3 Python Scripts with a README file for reproducibility
- Experimentation Recommendations
- Potential A/B testing experiments for the future