

User Segmentation and Product Recommendation

I. High-level approach

In this analysis, I employed 2 most popular clustering techniques to segment users, and then came up with recommendations for product improvement and marketing strategy. I used the survey data as the main data source for machine learning models, and the usage data to enrich the data insights. This helped us to segment users right after users join the app and fill in the survey, offering them customized experience as soon as possible.

After running 2 clustering methods (K-means clustering and Hierarchical Clustering) and examining the insights of clusters, I realized that the insights of segments produced by 2 methods are not much different. I decided to go with the output of K-means clustering, thanks to its more actionable insights!

II. Summary of user segments

Overall, subscription and primary language efficiency are 2 key features that help differentiate users. Users usually have better learning engagement if they are subscribers or/and know the primary language before using Duolingo. Besides, it is also interesting that Duolingo is better attracted by adult/senior learners than young people.

1. Highly-engaged subscribers

The segment accounts for 16% of all users. More than 80% are 35+ years old (50% are more than 55+ year old) with good income. They know the primary language before learning on Duolingo. With high commitment, they become the most engaged group.

2. Beginner Subscription Learners

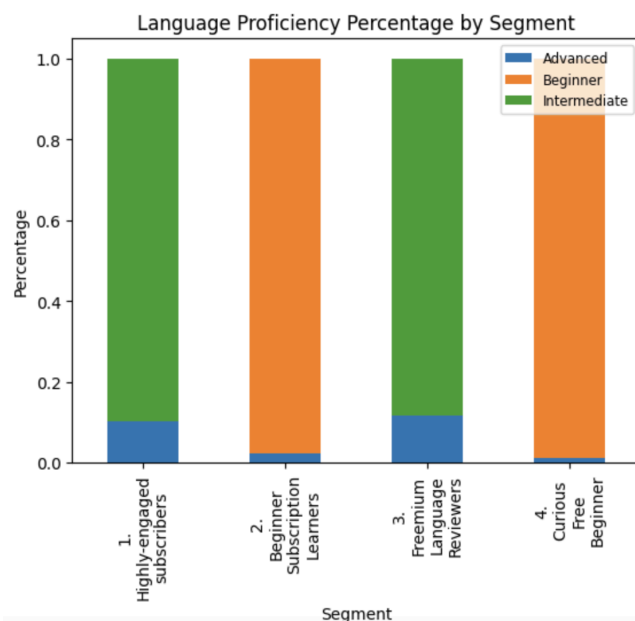
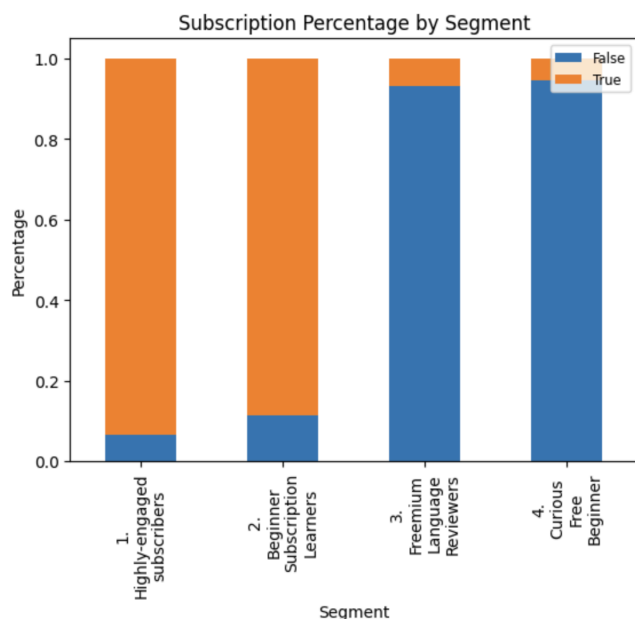
The group makes up 18% of all users. 80% of them are 35+ years old. Although they learn the language for the first time, their engagement and consistency is impressive.

3. Freemium Language Reviewers

This group accounts for 30% of all users. They are diverse in terms of age, income and employment status. They use Duolingo mainly to review languages they already know. Their engagement is medium - around 50% of subscribers'.

4. Curious Free Beginners

This is the largest group (36% of all users). They are just language beginners and want to explore the language. However, they seem stuck with lowest engagement. They also do not have a strong motivation/commitment.



III. Product and Marketing Recommendations and Next Steps

1. **Subscribers are our top prioritized segment** because they very likely generate the most revenue. The goal is to provide them with the best learning experience and make them re-new subscription. With high contribution of adult/senior learners (35+ years old), together with high commitment and engagement, we should:
 - Design learning paths with a focus on older adult learning styles and preferences
 - Offer more advanced courses challenges for the advanced learners.
 - Promote social features (such as Add friends) within the app to connect them with other engaged learners.
 - For marketing, showcase stories of successful late-in-life language learners and partner with organizations/companies catering to mature/senior audiences.
2. **For Freemium Language Reviewers, the goal is to convert them to subscribers**, because they are the potential group to become the Highly-engaged subscribers: have intermediate proficiency for primary language before using the app and medium engagement.
 - Highlight the added subscription values. Recognize their hard work and offer 4-week free trial subscription instead of 2 week ones
 - Introduce more advanced courses based on their interest to increase their engagement.
 - For marketing, emphasize the importance of reviewing a language, leveraging that precious foundation to continue improving.
 - As mentioned, they are a diverse group, therefore, we should further segment/analyze them to understand their challenges/pain points for more relevant strategy.
3. **For Curious Beginners, they are the biggest but riskiest group!** Most of them are beginners but they have very low engagement, potentially suggesting that they do not find it interesting to learn on Duolingo.
 - Quickly captivating the interest of this group at the beginning is important! Revamp the first courses to make sure that they can realize the quick win.
 - Offer personalized learning paths based on their interests and goals to increase their engagement.
 - Further analysis on their commitment/motivation is recommended to increase their motivation. The current data is not enough to tell why they have low commitment.
 - If watching ads is a reason leading to their low engagement, offering them 3-week free trial subscription and recognizing their improved performance after the trial.
4. **Next steps:**
 - Further analysis on the reasons why users subscribe and users' pain points are highly recommended to have a better understanding.
 - Conduct the user segmentation by country, understanding the reason why, for some countries such as Japan and France, the subscription ratio is much higher than others' and allowing to localize the product in each country.