Subscriber Optimization for Rosetta Stone

Joseph Calise, Ronan Kearns, Charlotte Belke, Nella Khachian, Ryan King, Josh Drake

Professor Houldsworth

MGSC 410-01, Group 5

November 30, 2023

Table of Contents

GROUP BREAKDOWN

Name	**Major/Minor**	**Skill Sets**	**Division of Work/ Project Contribution**
Joseph Calise	Computer Science Minor Analytics	Python, Scripting, Modeling	Action Plan, Modeling, Data Cleaning, Data Visualizations, Analysis, Slide creation, Document Writing
Ronan Kearns	Data Science Minor Chinese	Python, Modeling, Visualizations	Modeling, Data Cleaning, Data Visualizations, Document Writing
Nella Khachian	Business Marketing Minor Data Analytics	Tableau, Excel, basic Python and SQL	Tableau data visualizations, interpret findings, formatting, Slide creation, presentation
Charlotte Belke	Business Entrepreneurship/ Data Analytics	Tableau, Excel, basic Python, R, Visualization	Tableau data visualizations and interpretations, formatting, slide, presentation
Ryan King	Data Science Minor Game Development	Python, R, Visualization	Presentation, interpretations, and data cleaning
Josh Drake	Business Finance Minor Analytics	R, Python, Excel, Powerpoint, Presentation	Presentation, data cleaning, executive summary

INTRODUCTION

As the analytical team for Rosetta Stone, we analyzed both the app activity and subscriber datasets to gain insights and extract significant patterns and trends. Our analysis goal is to pinpoint the premier subscribers within the existing data segmentation, to recognize those with the potential of being sold more products and services, and to determine the subscribers who are disengaging from the product. Our team will review a comprehensive analytical strategy, data cleaning and transformation, and exploration, which our data visualizations will complement to show our evaluation. In addition, we will target and explain procedures for specific business opportunities for Rosetta Stone to consider in the future.

ACTION PLAN

Given our five objectives for this project, we have devised an early, variable action plan to meet all objectives for this project efficiently.

1. ***Open Data:*** We plan to get into the data and familiarize ourselves with the two data sets provided and the data key provided in the final project folder. Utilizing the data key will provide the first steps in understanding the data we observe and the nuances in each set.
2. ***Data Discovery:*** Utilizing tools such as Excel, Pandas, and Tableau, we plan to execute simple functions to understand the robustness of our data sets, where the strengths and weaknesses of each data set, as well as grasp a consensus of how we can use this data. During this step, we plan to extrapolate information, including maxes, mins and means of variable columns, as well as understand the volume of missing data we are presented with.
3. ***Data Cleaning / Transformation:*** Utilizing the same tools as above, we aim to clean each data set accordingly, without the exclusion of a large amount of rows. In fields of numbers or averages, we plan to fill in as much missing data as possible to preserve that row of information for later analysis. During this step, we will also transform data and create columns for better analysis and modeling techniques later in this project.
4. ***Early Visualization:*** During this step, the group will use the newly cleaned data to create robust visualizations for early analysis with hopes of grasping early insights from the visualizations created.
5. ***Modeling:*** With our cleaned data, we will train and test various models we see necessary to obtain more insights to answer the questions provided and meet the objectives. We will note all insights we have found for later consolidation and closer analysis during this time.
6. ***Consolidation and Analysis:*** We will take all information from the models in step 5 and correlate the results to our project objectives to drive further insights and understanding of the outstanding questions.
7. ***Documentation:*** Throughout the work, we will be documenting as much information as we can to provide in a handwritten format alongside the final presentation. We will then

use this documentation to create our final presentation utilizing the insights gained during the above steps.

8. ***Presentation Creation:*** We will create our final presentation based on the information in the above steps to answer each question laid out for the final project.
9. ***Presentation Practice:*** We will practice our presentation to ensure proper information delivery during our final presentation.

ABOUT ROSETTA STONE

Rosetta Stone provides diverse pricing structures tailored to the number of languages one desires to learn, the duration to finish a course, the chosen platform, and the preference for personalized coaching. To embrace its mission to transform lives through language education, Rosetta Stone strives for widespread accessibility, fostering education and profitability. Their approach involves streamlining software accessibility, incorporating comprehensive business-related content, leveraging speech recognition technology, and providing unlimited live tutoring. To ensure uninterrupted service, Rosetta Stone has employed consistent LMS (Learning Management System) integration, allowing for more seamless exchange and update of data within existing technologies. They also offer different analytics for tracking employee engagement and gauging return on investment. Their flexible licensing not only simplifies the sharing of resources but also bolsters enterprise-level support, simplifying the tool usage for those who aren't IT experts.

For marketing outreach, Rosetta Stone utilizes Google's TrueView advertising model, which is a cost-effective approach where payment is based on actual ad views. Their promotional content is strategically placed on Youtube and various applications to maximize visibility among potential users, targeting the platforms where prospective learners are most active.

DATA EXPLORATION

- Data Dictionary: This sheet explains various fields that are in the Subscriber Information and App Activity datasets to help bring clarity during analysis.
- App Activity: This dataset is for App Activity, and in the file it has ID, App Session Platform, App Activity Type, and App Session Date
    - **ID**: identification for session of app activity
    - **App Session Platform:** The platform in which was used to access the content (IOS or Android)
    - **App Activity Type:** Actions taken by user: App launch, Start (lesson or unit), Completed (lesson or unit), Onboarding content, or Other (any other action taken in-app)
    - **App Session Date:** The date of subscriber activity

- Subscriber Information: This dataset contains information that relates to the subscriber. It has categorical (e.g., language) and summarized data (e.g., email engagement); email metrics reflect the 90-day period ending 3/31/2020. Here are the specific fields:
    - **ID**: identification for subscriber information
    - **Language**: Language selected by subscriber
    - **Subscription Type:** Limited means the subscription has a defined expiration, usually 3, 6, 12, or 24 months from the start date. Lifetime mean it is a perpetual subscription, indicated by end date in year 2098 or 2099
    - **Subscription Event Type:** The type of purchase either initial purchase or renewal
    - **Purchase Store:** Web or App
    - **Purchase Amount:** The amount paid for subscription, available for only web purchases
    - **Currency:** The currency for the purchase amount
    - **Subscription Start Date:** The date that the subscription begins
    - **Subscription Expiration:** The date that the subscription ends
    - **Demo User:** The subscriber has used demo content on the app
    - **Free Trial User:** The subscriber registered for a limited time free trial (usually 3 days)
    - **Free Trial Start Date:** The date the trial subscription begins
    - **Free Trial Expiration:** The date the trial subscription ends
    - **Auto Renew:** Flag to indicate if the subscriber has turned off the auto renew option
    - **Country:** The country where the subscriber lives (self-reported)
    - **User Type:** Consumer or other (homeschool, business)
    - **Lead Platform:** Platform subscriber used to engage with Rosetta Stone products (web, app, or unknown)
    - **Email Subscriber:** The subscriber is opted in to receive emails
    - **Push Notifications:** The subscriber is opted in to receive push notifications
    - **Send Count:** The number of emails sent to subscriber in past 90 days
    - **Open Count:** The total number of times emails were opened by the subscriber in past 90 days (of those sent during same time period)
    - **Click Count:** The total number of times emails were clicked by the subscriber in past 90 days (of those sent during same time period)
    - **Unique Open Count:** The unique number emails opened by subscriber in past 90 days (of those sent during same time period)
    - **Unique Click Count:** The unique number emails clicked by the subscriber in past 90 days (of those sent during same time period)

DATA ASSESSMENT

The folder has three different data sets that are covered above. Each dataset contained unique information and was organized in a way that all of the columns were named and labeled. When scanning through the datasets in excel it was clear there were many ‘NA’ and ‘’Null’ values in the Subscriber Info and App Activity datasets.

Subscriber Data:

When accessing the subscriber information file provided, during our data discovery we have columns that prove to be problematic when it comes to NULL values. We noticed that there are specific categories where we tend to find the majority of our NULL values. They are purchase price and currency, free trial dates (start and finish) as well as our columns surrounding notifications and clicks. Below is a screenshot of NON-NULL value counts in each category as well as our data types and each column listed.

```
Data columns (total 24 columns):
 #   Column                   Non-Null Count  Dtype
---  ------                   --------------  -----
 0   ID                       40102 non-null  int64
 1   Language                 40102 non-null  object
 2   Subscription Type        40102 non-null  object
 3   Subscription Event Type  40102 non-null  object
 4   Purchase Store           40102 non-null  object
 5   Purchase Amount          26923 non-null  float64
 6   Currency                 26924 non-null  object
 7   Subscription Start Date  40102 non-null  object
 8   Subscription Expiration  40102 non-null  object
 9   Demo User                40102 non-null  object
 10  Free Trial User          40102 non-null  object
 11  Free Trial Start Date    5833 non-null   object
 12  Free Trial Expiration    5833 non-null   object
 13  Auto Renew               40101 non-null  object
 14  Country                  40102 non-null  object
 15  User Type                40102 non-null  object
 16  Lead Platform            40102 non-null  object
 17  Email Subscriber         40102 non-null  object
 18  Push Notifications       40102 non-null  object
 19  Send Count               28448 non-null  float64
 20  Open Count               28448 non-null  float64
 21  Click Count              28448 non-null  float64
 22  Unique Open Count        28448 non-null  float64
 23  Unique Click Count       28448 non-null  float64
```

Figure 1.

DATA CLEANING:

App Activity Data

The app activity data included the user actions within the app. This data included the platform they used, the action they took, and the date. To assess and clean the activity data, we used Python to check for missing or null values. Nearly half of the activity data, 44718 actions, had no platform information. Since these activity types for these actions were necessary, we filled the empty platform entries with ‘Unknown.’ Once the platform was filled, there were

14420 entries without an Activity Type or Date, which wouldn't be helpful for our data. Therefore, we removed these remaining 14420 rows so that the data only included complete data. Prior to cleaning the activity data had 809478 data points, which was cleaned down to 795058 rows, retaining the most amount of useful activity data possible.

Subscriber Data

In order to clean a data file like the subscriber data, the team needed a working plan to approach three main issues we observed with our data. Most surrounding NULL values, we needed to figure out how to work with or replace NULL values in Purchase Amounts and Currency, Free Trial Dates (start and end), and NULL values in the notification section located at the bottom of the figure above. We elected to handle each of these columns uniquely as it is not viable to drop all NULL values in the dataset, this would incur a loss of over 85% of our data.

<u>Purchase Amount and Currency:</u>

After reading through our data dictionary, we understand that the majority of NULL values in these columns occur when the Purchase Store is an App and not the Web. Understanding this, **the team felt that the Web purchase data was the most reliable to move forward with when calculating averages that will be used later in this section.** This idea did not come without its own issues. After sorting for Web purchases only in our data, we observed that our averages and data were being skewed heavily by about 1,500 rows of purchase amounts equating to 0. These rows of 0 had no commonality, and to the best of our knowledge were not a correct metric to weigh averages. The team believed that replacing these 0s would give us the most accurate depiction of our web purchases as a whole.

The team developed a script that combed through every row in the dataset and **first set every single currency to the accurate value in USD, based on conversion rates observed in 2020.** The team was then able to group the data and pull the average amount paid for every single language observed as well as the subscription type for that language.

After storing those values in a dictionary, another script was developed to replace each of those 1,500 rows that show the purchase amount as 0 and replace it with the average according to that row's language and subscription type. The cleaning techniques from above yielded the percentile results below for our Purchase Amount / Currency:

```
mean         69.157146
std          77.339979
min           0.000000
25%           0.000000
50%          35.970000
75%         119.000000
max         996.030000
Name: Purchase Amount,
```

Figure 2.

```
mean         88.137934
std          68.638928
min           1.090000
25%          39.000000
50%          61.650000
75%         119.000000
max         996.030000
Name: Purchase Amount
```

Figure 3.

Click and Notification Data:

When dealing with all the columns that were encapsulated in the click and notification data, the team believed that while there are a large number of NULL values, leaving these columns intact was the best way to preserve as much data as possible. The reasoning behind this decision was due to the fact that any modeling we will conduct utilizing these specific columns, will be run with the non-null observations we have. Instead of dropping almost half of our data because of these columns, we would perform that action in a control dataset only used for modeling with those columns.

Free Trial Dates:

The observation of the most NULL values falls in the Free Trial Start and End columns in our subscriber data. The team did not find these columns necessary to clean due to the fact that we have a column that will notify us if these columns will contain NULL. That column being Free Trial User, which if Yes, then we know we will have values in our start and end date columns. If not, then we know that these columns will be NULL and can handle these accordingly. We only intend to use this data for observations on the effect of the use of the Free Trial System and will only model and visualize based on those users who did in fact use the trial.

Extra Datasets:

After all the cleaning mentioned above, the team felt that the use of other datasets would be useful in certain modeling and visualizations we planned to create. We created two more notable datasets, one dataset named **appFilled.csv** and another named **merged.csv**.

appFilled.csv:

The idea behind the creation of appFilled was to replace the unreliable app purchase amounts with averages of the observed language and subscription type. The team believed this would open up a significant amount of visualization behind the total revenues and purchases being able to utilize all purchases observed, not just the web. The team was careful not to utilize appFilled for any purchase amount modeling due to the nature of the model being able to observe that there are averages and match the values to the patterned.

Merged.csv:

The idea behind creating a merged dataset was to be able to encapsulate the data from both subscriber data as well as the app activity data. This dataset was created by observing the number of interactions a user had in the application activity data and created columns in the subscriber data for each interaction tracked. The team found this dataset useful specifically in creating models for predicting whether auto renew is on or off for a user. We found the merging of this data to be necessary to observe the most important attributes for auto renew predictions, which will generate insights for #4 of our presentations.

EXPLANATIONS/FINDINGS

1. *Determine the most valuable subscribers*

When identifying our most valuable customers, we found it vital to define what we believe to be "valuable." The easiest metric we found to define this is simply revenue contribution, i.e. who are the customers that bring in the most money to the company, and why that is. Through our research to define our most valuable customer, the team stumbled upon a caveat to consider when deciding to increase the number of customers in this "most valuable customer" category. Within our business model, increasing the number of our most valuable customers can adversely affect the business and prove to be detrimental in the long term if other factors are ignored.

Research:

To answer any questions based on the data, the team found it necessary to create insightful models and visualizations to begin finding answers.

Models:

With our primary focus being revenue, the group attempted to create a linear regression model that would use all available variables to predict a purchase amount for that customer. Since the web data is the most reliable regarding purchase price, we elected only to use that subset of data to train and test the model. If we were to use the data we included averages in, the model would simply pick up on the pattern and will likely prioritize other attributes for its predictions.

Since the group was never concerned with predicting a purchase amount, the model was intended to be consistently used to identify important metrics for higher purchase prices. We found this model to be reasonably accurate in predicting. Still, regarding the information on the variable importance of the model, we didn't excavate the insights we hoped for since the variable importances were dominated by lifetime vs. limit subscriptions. This understanding drove the idea behind our investigation between upfront customer value vs. customer value left.

Visualizations:

Without adequate information in the model, we resorted to visualization methods to gain a deeper understanding of things such as our most popular languages, languages with the longest average subscription duration, and overall revenue generation from our products.

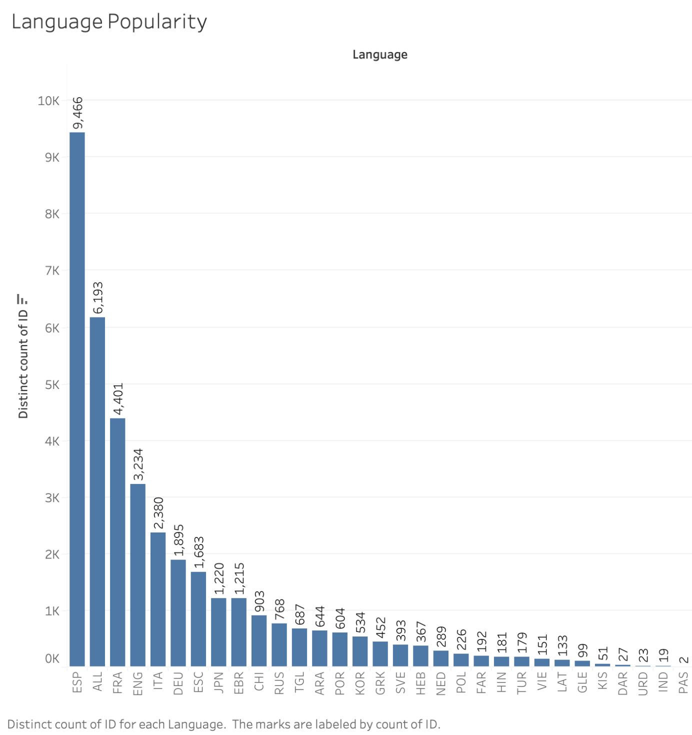


Figure 4.

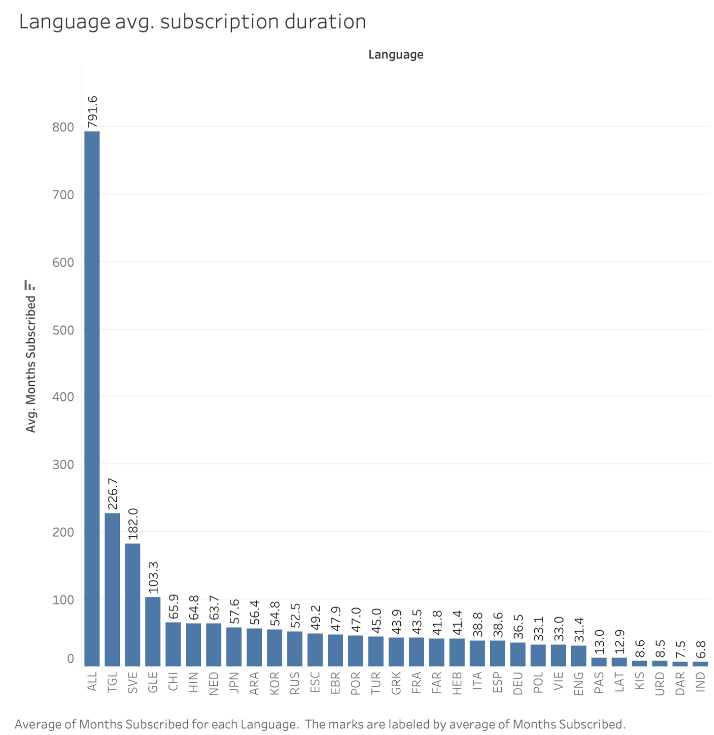


Figure 5.

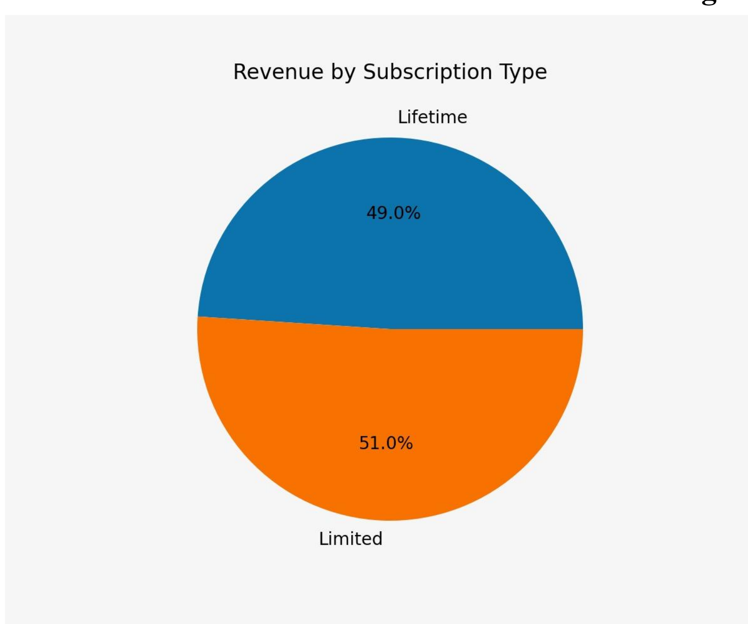
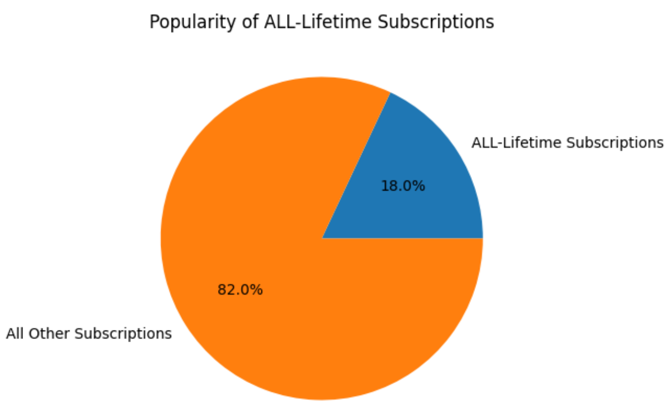
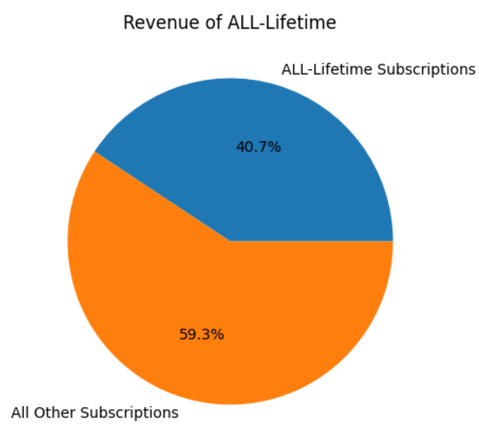


Figure 6.

Findings:

Most Valuable Customer (ALL-Lifetime): When looking strictly at revenue generated by a customer, our most valuable customer is those who subscribes to all languages for a lifetime. **These customers alone equate to 40.7% (web purchases only) and 47.9% (app and web purchase) within our observed revenue in our dataset .** With a definition of value being strictly the revenue a customer is bringing in, these customers are by far the most valuable to the company. To visualize this idea, we created two graphs below, the one on the left shows the amount of customers who purchased ALL-Lifetime subscriptions vs. every other subscription available and on the right, we posted the revenue generated between those two groups.





The visualization above shows clearly that while the number of ALL-Lifetime subscriptions only make up 18% of our subscription sales, it drives almost 41% of our revenue alone. **This only reinforces that ALL-Lifetime customers are by far the most valuable we have in our business.** The idea of creating more ALL-Lifetime customers, will certainly drive revenue, but for long term success of the business it can prove to be detrimental. As mentioned above, in this business model, it is not as easy as increasing the number of customers in the most valuable category. **If this business utilized every technique to enroll customers into ALL-Lifetime subscriptions, the business would drive revenue right away, but would lose further revenue from those customers. When a customer subscribes to an ALL-Lifetime subscription, there are no further products that can be sold to that customer, thus losing any future revenue potential. The group believes that prior to making attempts to increase the number of customers subscribing to the ALL-Lifetime plan, the business will need to weigh the customer acquisition costs vs. the revenue generated by converting customers to ALL-Lifetime. This is due to the fact that if we convert an existing customer to ALL-Lifetime, we can no longer sell them additional products, meaning we will need to continue acquiring new customers.**

2. Understanding the subscriber segments present in the database

When segmenting our customer base, the team did a more profound analysis than just the typical segmentations like geographical or other standard means. To do so, we created various models and visualizations to further divide and understand our customer segmentations. Ultimately, the group believes they have defined customer segmentations and behaviors that help the company differentiate the behavior around each customer segment.

Modeling

Below is a description of every model we tried to further segment our customer base, with interpretations of each model.

Hierarchical Agglomerative Clustering

In order to comprehend the subscriber segments within our database, we employed Hierarchical Agglomerative Clustering (HAC) as a key analytical technique. HAC is a hierarchical clustering method that organizes data points into a tree-like structure, providing insights into the inherent structure and relationships within the dataset. The choice of HAC, specifically utilizing Ward's method, was driven by its ability to minimize the variance during cluster formation, which revealed distinct patterns in subscriber behavior.

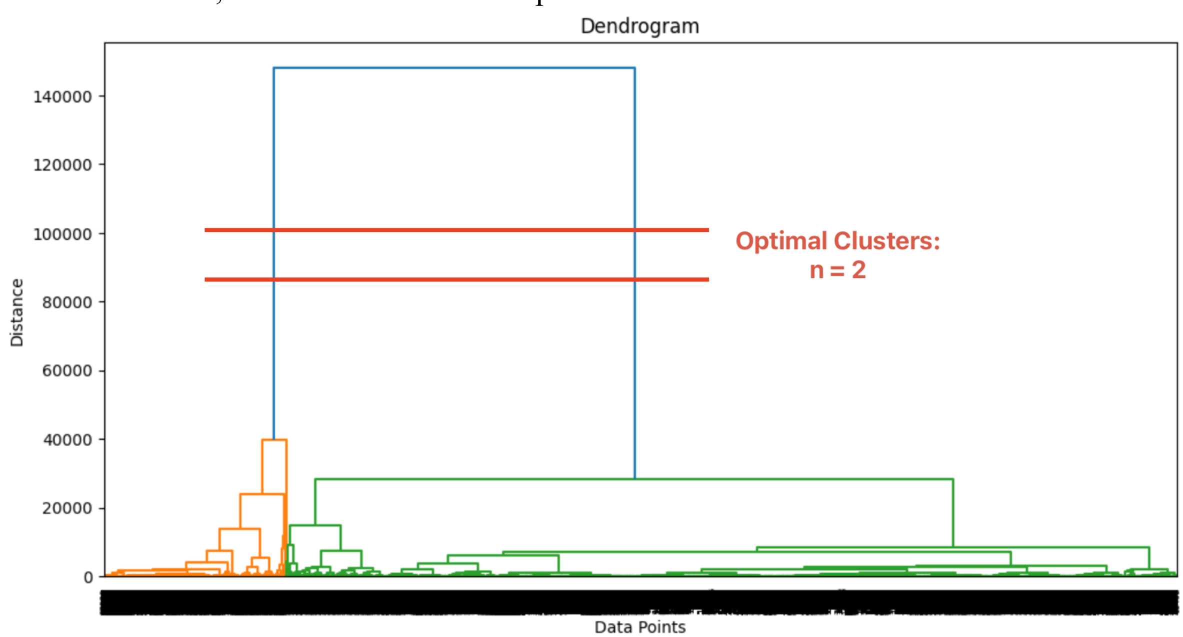
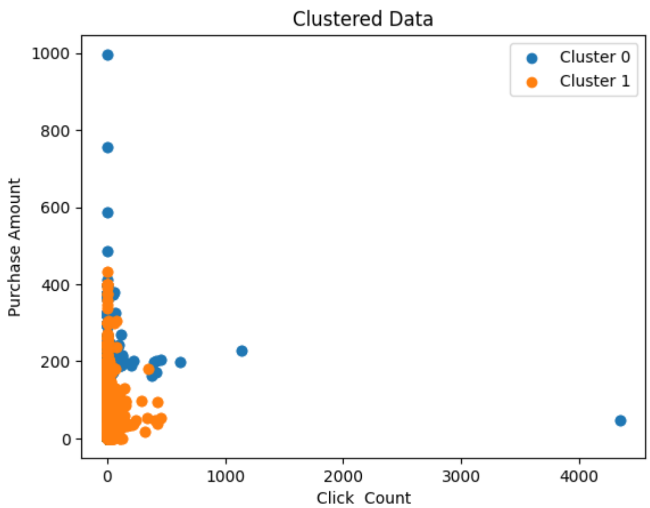


Figure 7.

To initiate the clustering process, we first prepared our dataset by handling missing values and encoding categorical variables. Subsequently, the linkage matrix derived from the HAC algorithm was used to construct a dendrogram, Figure 7 above, offering a visual representation of the hierarchical relationships and history of groupings among the subscribers. We identified two clusters as the optimal amount by finding the largest vertical distance that doesn't intersect with other clusters.

Following the dendrogram analysis, we applied Agglomerative Clustering to assign cluster labels to individual subscribers using two clusters. The cluster assignments were calculated by the inherent characteristics of our data shown in the dendrogram. This step enabled the classification of subscribers into two distinct segments, laying the foundation for a deeper understanding of subscribers' characteristics within each cluster. Included Figure 8, we graph these clusters to see the behavior of these clusters between purchase amount and click count.

Figure 8.



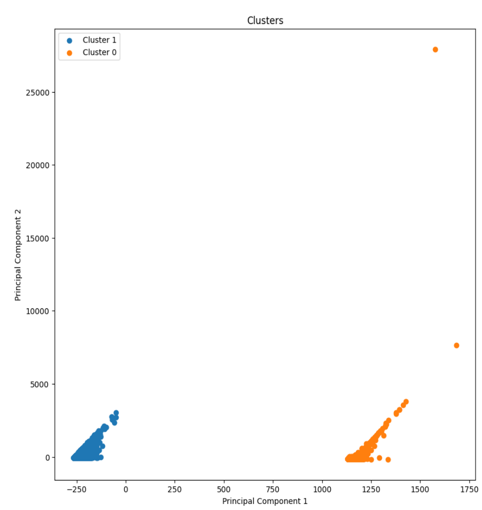
<u>Orange (Cluster 1): Price Motivated</u>
Subscriber's behavior in this cluster indicates they prioritize price over content. This is characterized by lower purchase amounts, indicating a emphasis on cost-effectiveness. While the clusters are slightly overlapping, there were more subscribers Clustered in Orange than Blue. This reveals that this cluster typically has more subscribers with higher click counts.

<u>Blue (Cluster 0): Content Motivated</u>
Cluster 0 contains subscribers whose data suggests priority towards long-term content over price. The blue clusters data points have a higher purchase amount but not as many subscribers with higher click counts. This reveals a willingness to invest in the language learning experience, prioritizing the value of the content over price considerations. These visualized clusters provide a tangible representation of the distinct behavioral patterns observed in the data.

Principal Component Analysis

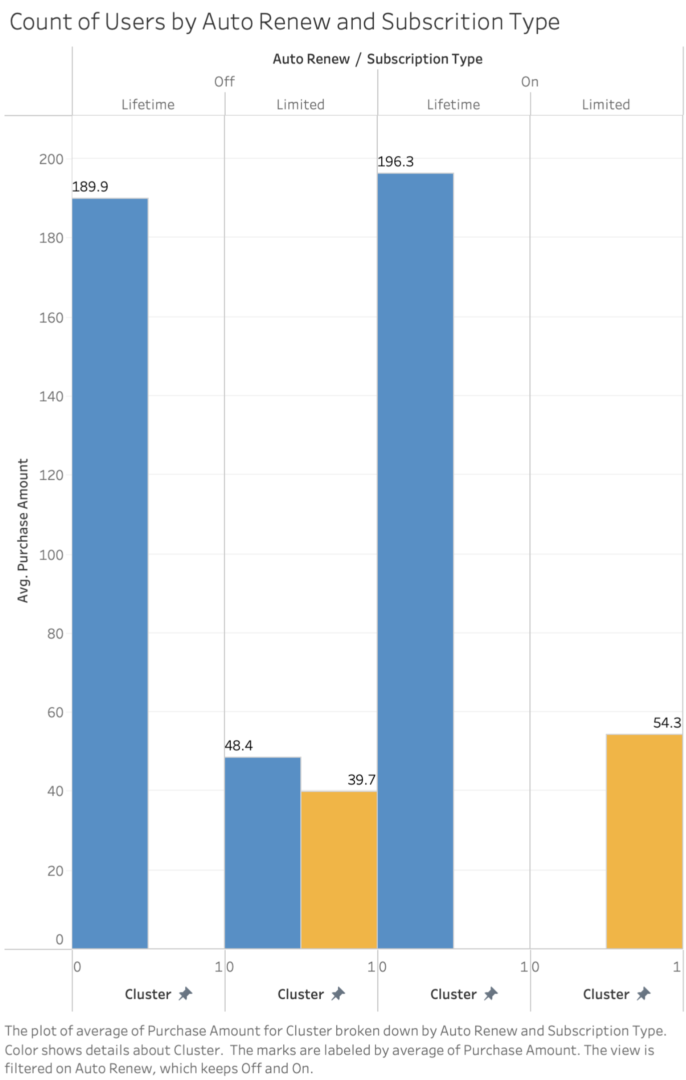
Our group performed Principal Component Analysis (PCA) to interpret and visualize the clusters further. By reducing the dimensionality of our dataset, PCA allowed us to represent subscriber clusters in a two-dimensional space. The resulting scatter plot effectively visualizes the relationships between clusters, aiding in identifying patterns and trends within each subscriber segment.

Figure 9.



After using HCA and PCA to uncover subscriber segments by cluster, we were able to understand these segments further with Tableau. With Tableau, we created visuals to compare important variables against each subscriber's assigned cluster to reveal characteristics unique to each cluster.

Figure 10.



In Figure 10, the bar chart shows the average subscribers purchase amount by whether they had auto-renew status of on or off. Furthermore, the averages for either status were separated by limited and lifetime subscriptions. **The color of the bars represents the assigned cluster of the subscribers who made up the bar's average.**

A majority of subscribers were grouped in Cluster 1, shown in orange. Subscribers in Cluster 1 had considerably lower average purchase amounts compared to Cluster 0 (Blue). This aligns with our findings from the scatter plot in Figure 9 which shows Cluster 1 having subscribers who were motivated by the cost of the subscription over the content they receive.

Turning focus to Cluster 0, the majority of subscribers had a significantly higher average purchase amount. This reflects the behavior of these subscribers being more motivated by the content they receive in the subscription. Furthermore, the two highest average bars for Cluster 0 had users with Lifetime subscriptions. The content motivation with the lifetime subscription shows that Cluster 0 users prioritize sustained access to the service and seek extended value over time with their high average subscription amounts. Unlike Cluster 1 (Orange) who prioritize cost-effectiveness, subscribers in Cluster 0 (Blue) appear to be more focused on the value derived from the content itself. The clustering and visuals above provided our group with a deeper understanding of subscribers, highlighting the differences in motivation between subscribers assigned to each cluster.

Results:

From our clustering models and visualizations, the team began to lightly define a behavioral segmentation to be combined geographical, platform and other segmentations. Based on our clustering models and interpretation, visualization and further analysis, we feel confident to define one more segmentation between a group we call content motivated customers and a group we call price motivated customers.

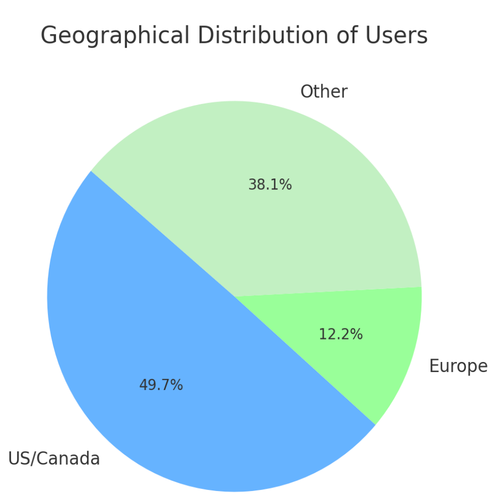
Below is the overview of our customer segmentation.

Behavioral Segmentation:

- **Customers who value learning the language over price (motivated learners), these customers display the behaviors of:**
    - Having auto-renew **on** more often
    - On average pay 10% more than their counterparts
    - Have more interactions with the application
    - Have more completed units and modules in the application
- **Customers who value the price they are paying more than learning the language, these customers display the behaviors of:**
    - Having auto renew **off** more often
    - Pay ~10% less on average
    - Have less overall interactions with the application
    - Have less complete units and modules in the application

We felt it important to include this behavioral segmentation, because while we do not know the reasons, we are able to extrapolate trends between these segmented groups and make assumptions from there. For example, **with the content motivated group, we believe these people may be taking a trip to a different country soon, or may have a deadline to learn a language for a job or something else happening in their life. They are willing to pay more to obtain the content right away and start learning, interacting with the application more and completing more modules.** On the other hand, **we can loosely identify price customers who value price over learning the language as people who may want to learn a language for fun. These people likely do not have a deadline to learn the language, and rather than obtaining the content right away for a higher price, they likely wait until there are deals and promotions for the content they are looking for.**

Figure 11



Geographical Segmentation: Our geographical segmentation includes what percentage of our customers are located each geographical category:

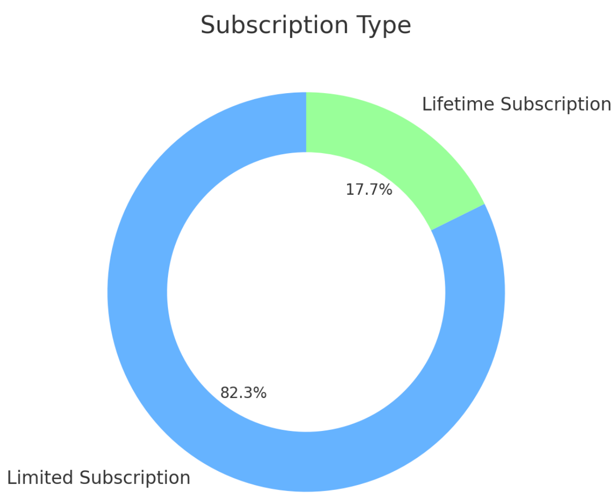
U.S/Canada: 49.7%

Europe: 12.2%

Other: 38.1%

A geographic segmentation is important in any business model to identify areas where the customer base is more concentrated. For example, for Rosetta Stone, we have a higher concentration in the U.S/Canada than Europe by a lot. If we were looking to diversify our customer base, this segmentation would allow us to quickly identify where we should focus our diversification. On the other hand, if we were looking to run a promotional campaign for existing customers, we may use this segmentation to identify that the promotion will target a larger amount of existing customers in the US/Canada region.

Figure 12

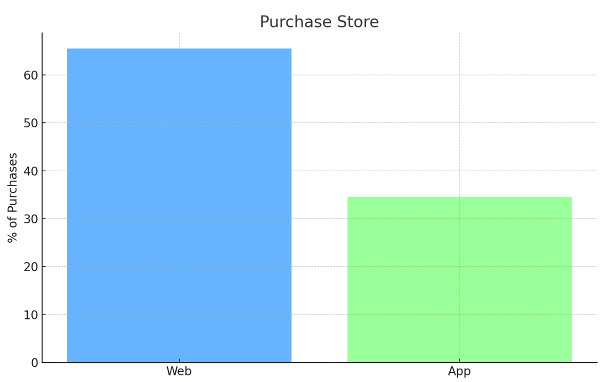


Subscription Type Segmentation: Our subscription type segmentation allows for us to identify the split amount Limited subscribers and Lifetime subscribers:

Limited Subscriptions: 83.1%
Lifetime Subscriptions: 17.9%

Using this segmentation, we can quickly identify the gap between both our subscription types offered at Rosetta Stone. This will allow cost value analysis for the company to perform to find insights into the disparity between the two subscription types.

Figure 13
Purchase Store Segmentation:



This segmentation allows us to identify popular purchase stores for our customer base:

Web: 65.5%
App: 34.5%

Utilizing this segmentation, we can identify possible trends between our customer's preferences between Web and App purchases. While this segmentation may not generate insights right away, this is a segmentation we can keep an eye on over time and identify any shifts and investigate the reasoning behind those shifts.

Overall, the above segmentations yield different insights into different aspects of Rosetta Stone's business. Segmentations like behavioral and geographic can drive insights and decisions right away, while segmentations like Purchase Store, may be a segmentation you can monitor and notice trends over time.

3. *Identify the most likely subscribers who could be sold additional products or services*

To find who we believe to be the customers most likely to be sold additional products, we felt that leaning heavily into customer's email behavior would be vital information to assess. The group values this idea the most because in order to sell additional products, we will need to rely on customer engagement via the emails Rosetta Stone sends out. If customers aren't receiving or opening the emails, that leaves a lower number of customers we can notify of additional products or deals.

Initially, the team considered running models to predict the number of emails opened by a particular user based on all other variables, but ultimately decided against it because of the significant amount of data to be dropped (> 80% of data loss).

We instead, created a new column in our dataset named "Open Percentage" and "Click Percentage". **Open Percentage** took the number of emails the customer opened and divided it by the number of emails we sent them. **Click Percent** was created by taking the number of times a customer clicked inside one of the opened emails and dividing it by the total number of emails that the customer opened. With these two columns created, the team believes we have the best metrics to analyze what customers not only open the most emails, but also those customers who <u>engage</u> with the emails the most by clicking within it.

When creating visualizations for these two new metrics, we isolated users who **open 25% or more of the emails we send them and have a 5% or more interaction rate with those emails they open.** We created this threshold based on benchmarks we found from **MailChimp, a popular email marketing company, which states in the market of education and training, the benchmarked open rate is 21%, while the benchmark click rate is 2.9%.** With this threshold, we are identifying customers who have above average open and click rates within our business. (https://mailchimp.com/resources/email-marketing-benchmarks/)

Unfortunately, these customers only make up **9.7% (3761 / 38611)** of our total customer base, and another 746 of those customers are currently **ALL-Lifetime** subscribers who currently have all the products for a lifetime, and have nothing else to purchase. **While the group believes this is an alarmingly low percentage of customers, this is the best metric to consider to identify customers whom we can sell further products to.**

We attempted to differentiate these customers further and grouped these customers by country and language to determine what customer base interacts the most with our emails and marketing based on the criteria set above.

First, the team wanted to know if there was a correlation between email interactions and the customer's location. We created a pie chart that clearly showed that the customers who interact most with our emails are in the <u>US/Canada</u>, followed by other countries, and then Europe. These percentages relate back modestly to our customer distribution by country, showing us there is not a large regional effect when it comes to customers who interact with our emails and marketing.

Figure 14

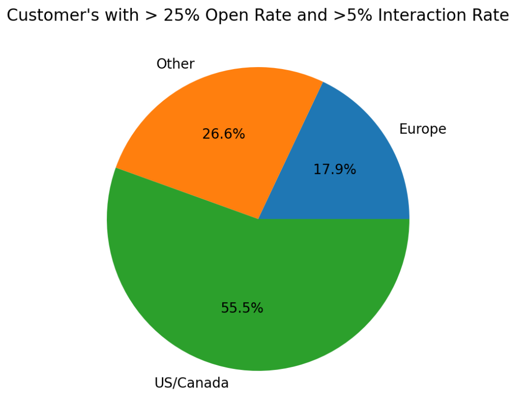
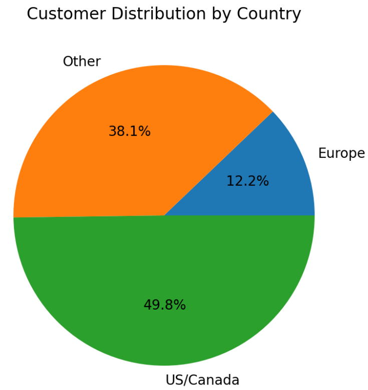
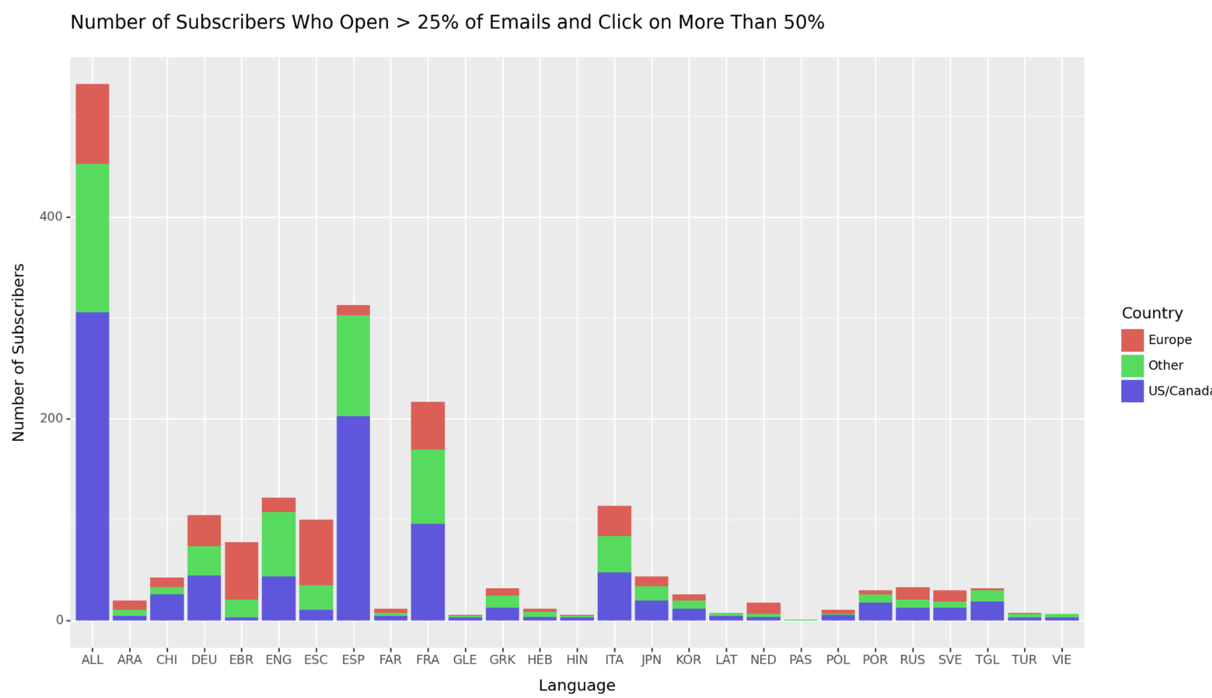


Figure 15



We then moved to create a graph that shows the number of customers who interact with our emails and marketing, grouped by language and country:

Figure 16



From this graph, we are easily able to identify the languages that typically have the most interactions as well as the country those customers are located in. We noticed a trend of our most popular languages and country origin within those languages for customers that meet our interaction threshold. **We notice above, the US/Canada and Other Countries tend to have more users interacting with emails when subscribing to ESP, ALL, FRA, and ITA (mostly European languages). While Europe has more customer interactions when learning languages like ESC, EBR, FRA, DEU, and ITA.** Breaking down these customers who interact according to our criteria gives us an insight into **possible products we can target towards these customers.** With customers from US/Canada and Other Countries a strategy may include promotions for other European languages or promotions for subscription upgrades to Lifetime. For European customers, we may change those promotions to other languages like EBR or ESC, while continuing the upgrade to Lifetime.

Lastly, to ensure accuracy and transparency, we included the ALL-Lifetime subscriptions in the customer data when analyzing interactions. These customers equated to **728 of our 3761** customers who meet our interaction criteria, we must note that while these customers interact with our marketing and emails, they no longer have additional they can be sold. For business purposes, these figures can be adjusted to a higher open and interaction percentage, but the team believed, based on the MailChimp benchmarks, a 25% open rate and a 5% interaction rate was the best metric to maintain as much data as possible to identify correlations, while still sorting the customers we believe fit.

Results of Customers to Sell Other Products To:

<u>We identified the group of customers who open 25% or more of their emails and interact with 5% or more of those emails they open</u> as customers who are likely to be sold additional products from our business. As mentioned above, the team leaned heavily into marketing and email interactions with our customers to identify this group. In a marketing and business sense, the best way to sell additional products to existing customers is through direct marketing, which, in our case is emails. While the team has identified these customers as most likely to be sold additional products, we can't help but to note that this customer base only makes up **9.7% of our total customer base within the dataset.** This is an alarmingly low number, and we believe that there should be a plan to address how the business can increase this value for further products and promotions.

Methods to Sell Additional Products:

From the analysis above, the team believes that we have not only found the group of customers most likely to buy additional products but also found an effective way to do so. We noticed two main ideas when referring to **Figures 14, 15, and 16**. The country in which our customers are located does not have a large effect on the rate at which they open and interact with our emails, and our customers that meet our interaction criteria, customers from different countries, are concentrated in certain languages. With this analysis, we can not only identify customers we believe will buy more products but also possibly target those customers with certain products or upgrades based on their country. From above, "We notice that above, the US/Canada and Other Countries tend to have more users interacting with emails when subscribing to ESP, ALL, FRA, and ITA (mostly European languages). While Europe has more customer interactions when learning languages like ESC, EBR, FRA, DEU, and ITA." Using this observation, we can target our marketing more effectively and offer promotions or additional products for specific countries based on the observations above. For example, suppose we are dealing with a customer who meets our interaction criteria, and they are located in the U.S., Canada, and other countries. In that case, it may be in our best interest to offer a promotion or additional products aimed at other European languages they do not have, like ESP, FRA, FRA, and ITA. On the other hand, we can do the same with European customers who meet our interaction criteria.

4. *Identify the subscriber profile of those not continuing with their usage of the product and identify the barriers to deeper subscriber engagement where possible.*

With the data provided, the team found that the easiest method to profile customers who are continuing their subscriptions from those who are not was by observing the Auto Renew column value in the provided data. After sorting by Limited subscriptions (since Lifetime subscribers do not renew), we noticed that of the 32,066 customers who have a limited subscription, 19,244 of those customers have auto renew off. This equates to 60% of our limited subscribers do not currently intend to renew their subscriptions. When it came to building a profile of this group of customers, we found it useful to find the differences between the two groups and from there generate a customer profile for those 19,244 customers not renewing. The team began this process by creating various models, some proving to be useful and most proving to be unsuccessful. We began with the intention of clustering this data to see if we could find the differences between these two groups. Unfortunately, the clusters of data in various different plots were too closely related not allowing for our algorithms to classify accurately. Next, we created promising logistic regression models and random forest classification models to accurately predict whether a customer will have auto-renew turned on or off. While these models performed well, we found it difficult to identify the positive and negative effects of the variables for the final outcome. But from the importance of variables these models weighed, we formed the idea to group and compare these customer groups using visuals and different grouping factors. With these techniques the team believes they have found the most important attributes that differentiate these two groups of customers and build a profile of those not continuing their subscriptions, and identify multiple barriers of engagement for these customers. It is important to note that in the observations below, the group observed only Limited subscription customers to identify differences between customers with auto-renew and without it.

Building a Profile

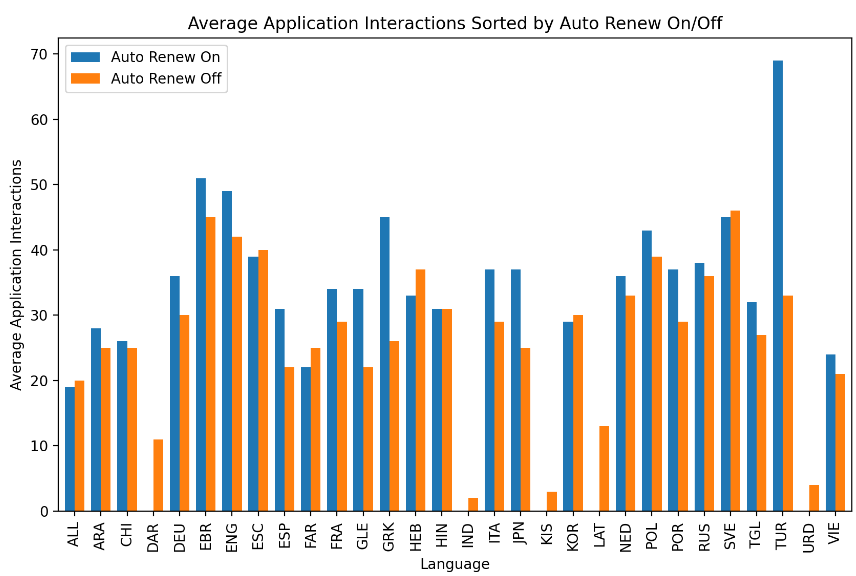
Customers with Auto Renew On Pay More: Our first group was to look at the relationship between Auto Renew and Purchase Amount. Our initial thought going into this grouping is that we would likely see that customers with Auto Renew on, likely are paying less, and that is what entices them to keep it on. To our surprise, the results were the opposite. We observed in both our cleaned data set only observing Web purchases, as well as in our merged dataset dealing with app and web purchases, customers with Auto Renew on pay an average of 10% more than those who have it off. While the team was trying to reason why this could be, we came up with a hypothesis that this points towards a group of customers (those with auto-renew on) who prioritize learning a language rather than the price they are paying to learn said language, we lightly called this group "motivated learners". On the other hand, this may say the opposite of the other customer group (those with auto-renew off), but the team wanted more evidence of that before defining that profile.

Customers with Auto Renew On Tend to Subscribe Longer:

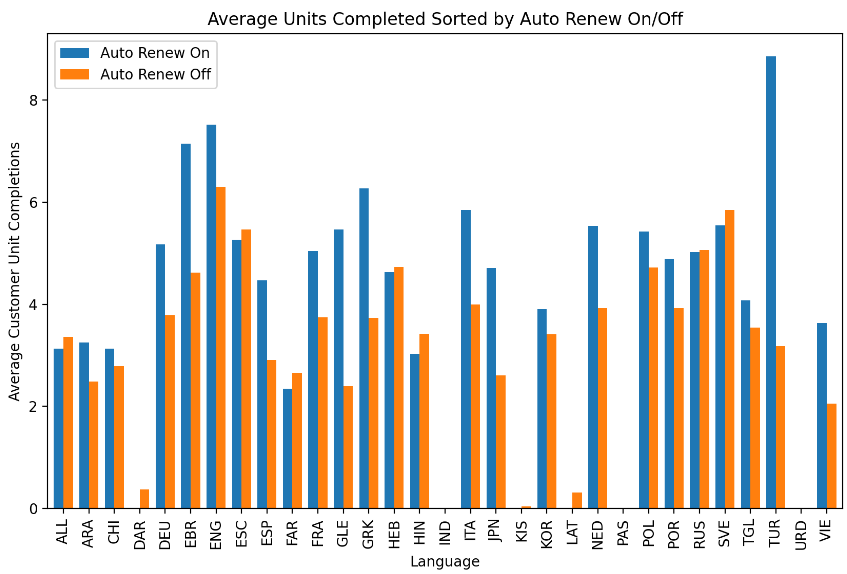
In order to build the "motivated learner" theory more, we wanted to compare these groups' behaviors further. When grouping the data, we found that between these two groups, customers who have auto-renew enabled pay more on average. We also found they normally subscribe for three more months than their counterparts. From our dataset, we found that customers who have auto-renew disabled subscribed for an average of **7.12 months**, while their counterparts subscribed for an average of **10.01 months**. Combining these two observations were intriguing, counterintuitive, and aligned with our hypothesis of motivated learners. It is important to note that while there is a correlation between paying more and subscribing for longer, the team checked every popular subscription length, and we constantly observed auto-renew customers pay more.

Auto Renew Customers Interact with the App More:

Lastly, we aimed to quantify this hypothesis within our data. In order to do so, we grouped each of the languages and then created sub-groupings of what we heard the customer had auto-renewed on or off. Utilizing Python and Pandas we obtained the average number of interactions each sub-group of each language had with the application. Below is a graph of those results.



From the graph above, the team can lean into this idea of a "motivated learner." We have first established that customers with auto-renew on tend to pay ~10% more and subscribe for ~3 months longer than those who have it off. Furthermore, this chart clearly shows that aside from outlying languages, customers with auto-renew on engage with the application more through various application interactions. We then took it one step further to solidify this motivated learner idea and wanted to see if these interactions carried over into units completed and found the same trend.



From these observations and figures, we believe we can establish two customer profiles through various metrics. We believe customers with Auto-renew can generally be classified as motivated learners; this group may include customers preparing to travel, needing to learn a new language for some aspect of their life, or someone who just prioritizes learning a new language. These customers generally value the content and learning more than the price they pay (within a reasonable amount).

On the other hand, we would not call customers with auto-renew off unmotivated learners, and we believe they fall more into the category of customers who see learning a language as less of a priority. These customers may be learning a language as a hobby or to fill their free time; they are less likely to have a "deadline" to fulfill and likely value a lower price for the content they receive.

With these metrics front and center, we believe there is further evidence that supports these two profiles of customers, inadvertently, as well as insight into possible barriers of engagement with the customers who are electing not to continue.

Identifying Barriers of Engagement:

Email Subscribers: When it comes to the number of customers who are subscribed to emails, our overall data equates to only **47.5%** of customers being subscribed to emails, and of those customers subscribed to emails, the average 'open' percentage is **~26%**. Based on

MailChimp benchmarks, our average open rate is slightly above standard, which is a positive sign. The group focused on the discrepancy between the two groups (Auto Renew On/Off) based on the percentage of customers subscribing to emails. When looking at customers with auto-renew off, we see only about 43% of that customer base is subscribed to emails, while 53% of customers with auto-renew on are subscribed. We have already defined that customers with auto-renew on typically engage with our application more often. While increasing this rate for both groups should be a priority, the group identifies that the 43% of customers with auto-renew off and not subscribed to emails remain a major barrier to engagement. **We identified this group to have lower levels of engagement with our application on average, as well as a substantial drop off in email subscription data. To drive deeper engagement with our business, we recommend a focus on this segment of customers.**

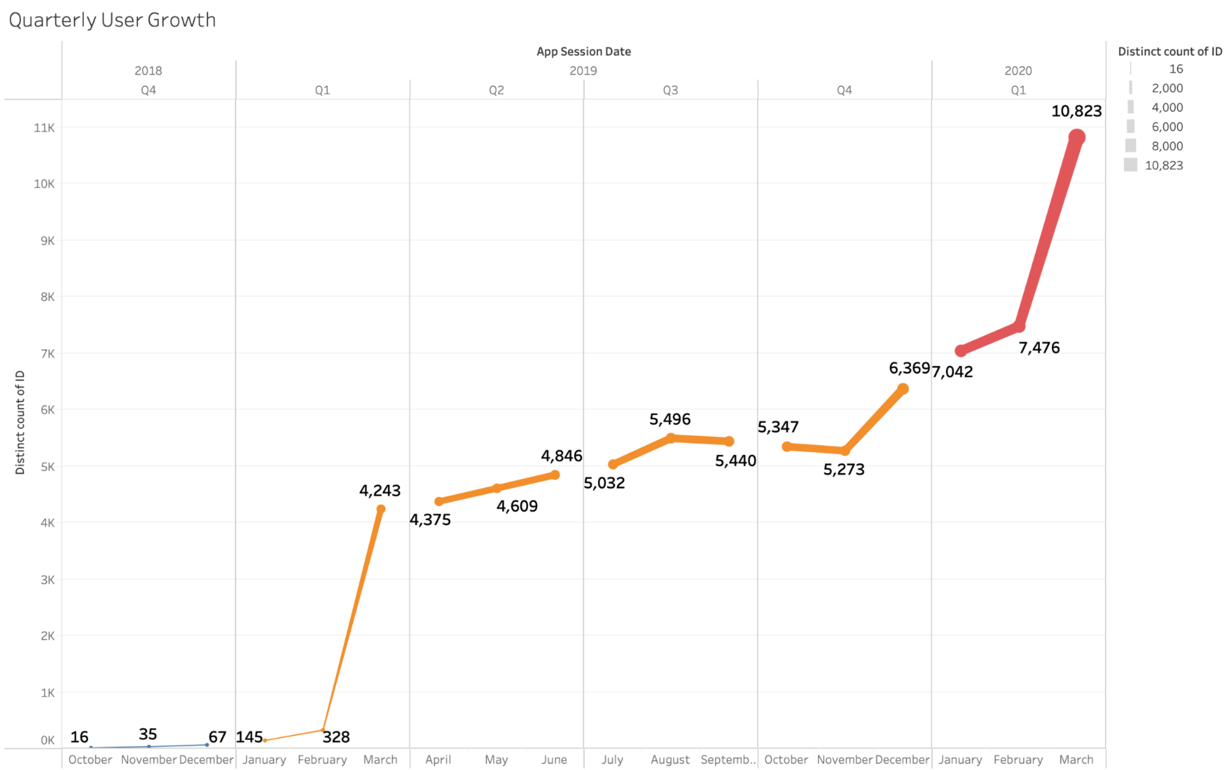
Marketing Techniques:

After establishing the discrepancy in email subscribers, the team found an intriguing similarity and difference between the two groups of customers. **On average, we send more than double the amount of emails to non-renewal customers than those who have auto-renew on. Between the two groups, on average, the non-renew customers open more emails than our renewing customers.** The team believes the reasoning behind this observation could be one of two things: non-renewing customers may be looking for deals before continuing their subscription, **or our emails being sent are not effective in driving engagement with our application and, in turn, are not converting non-renewing customers to customers who renew.** Since we believe that if non-renewing customers were looking for better deals, we would have a higher email subscription percentage, the team feels the latter idea is a more accurate assumption. With a higher opening count from our non-renewing customers, there is clearly a marketing and engagement gap via emails. While we have customers opening emails, we need to put ourselves in a better position to not only convert our non-renewing customers but also to drive engagement through our email marketing techniques.

5. ***Outline any business-relevant opportunities that are present from your analysis of the data not covered above.***

Our Most Popular Months:

During our data discovery phase, the group took some time to plot any information to understand the data at a deeper level. During this process, a group member created the graph below, which gave us an interesting insight we wanted to investigate further.

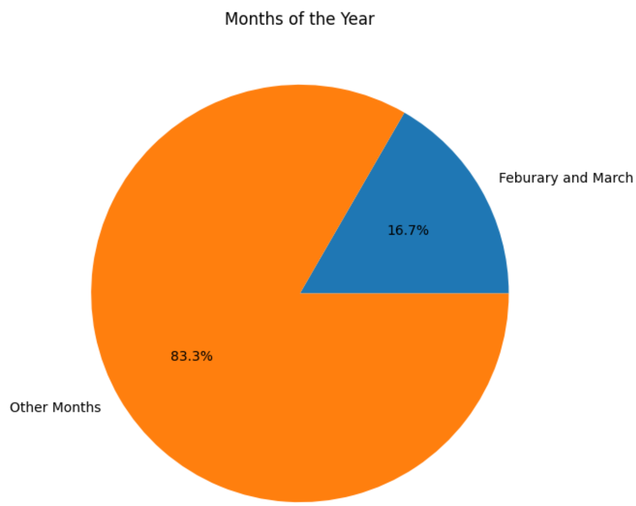


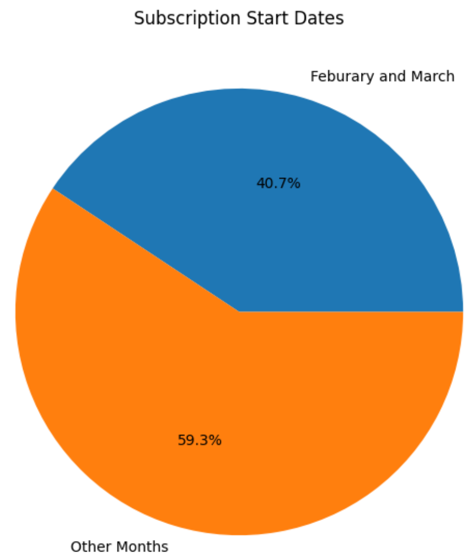
The trend of distinct count of ID for App Session Date Month broken down by App Session Date Year and App Session Date Quarter. Color shows details about App Session Date Year. Size shows distinct count of ID. The marks are labeled by distinct count of ID.

The graph above is created from the **app activity** dataset, and it plots distinct customer IDs quarterly over the time period of our data. **An interesting trend that popped out right away was the unusually high growth our activity experienced in the months of February and March.** In the short time we have captured in this dataset, the occurrence of is recreated for both observations of the end of Q1 for 2019 and 2020. The group wanted to further understand why this could be happening. Are people just on their applications more during these times, or are we seeing an increase in the number of subscribers we are generating in these specific time periods? To answer these questions, we took a series of steps. We first took our **subscriber** dataset and created a column to capture the month customers **began** their subscriptions. Utilizing Pandas and visualization, we then compared the number of customers starting a subscription during these months vs. the other months in the year. If our hypothesis is correct, we should see a significant amount of customers beginning their subscriptions during February and March.

Our Findings:

After forming our hypotheses, we set out to confirm them. First, we wanted to ensure that the application activity results from more customers during this period. After creating our Months column in the subscriber data, we utilized a pie chart to visualize how many subscribers start in the months of February and March, as opposed to other months.

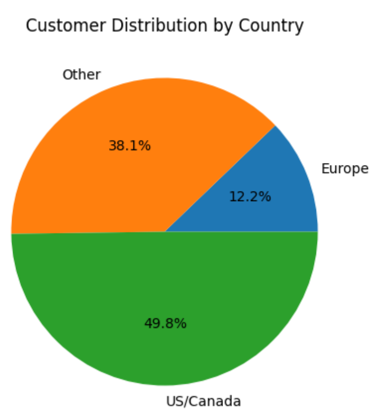


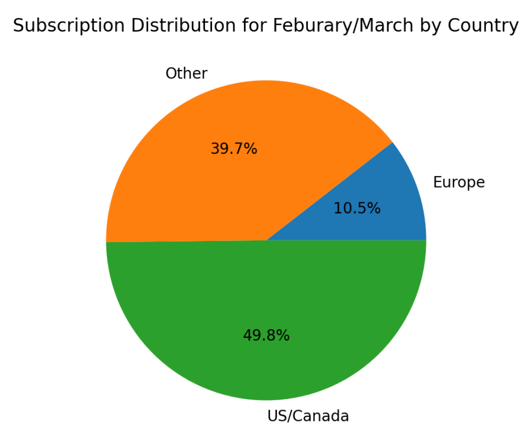


In the figure on the right-hand side, we found it useful to plot the distribution of these two months in terms of the year is useful. The figure on the right represents what our distribution would look like if every month of the year saw an equal amount of subscribers. The figure on the left visualizes the number of subscribers we gain during February and March as opposed to every other month in the year. **We can see that 40.7% of our customers subscribe during these two months alone.** The team believes this observation confirms our hypothesis, and we then turned to hypothesize why this may be the case with our customer base and to leverage this idea further for business purposes.

Further Analysis:

The team set out to find what insights or trends we can find within this subset of customers who subscribe between February and March. When creating visualizations, we again saw that location did not seem to have a major effect on the subscriptions starting in February or March.





Utilizing the visualization above, our subscription distribution closely aligns with our overall customer distribution, allowing us to conclude that geographic location likely does not play a significant role in this occurrence.

Possible Causes:

Summer is a popular time for travel, and people may be trying to learn languages during this time for their vacations. The increase in subscribers during February and March might also be due to seasonal factors. For example, these months could coincide with a New Year's resolutions push where people are motivated to learn new things, including subscribing to new services.

Results:

Rosetta Stone can increase marketing efforts during these peak subscription months to capitalize on the high conversion rate. Based on the insight that subscription increase is not heavily dependent on geography, we can run more global promotions during these months to incentivize customers. Offering bundles for popular destinations like France and Spain or activities people are interested in during February and March might attract more subscribers.

MAIN TAKEAWAYS:

Objective: Analyze Rosetta Stone's subscription base to identify valuable customer segments, focusing on customer behaviors, geographical distribution, and subscription types.

Key Findings:

Revenue Drivers:

ALL-Lifetime subscriptions, though only 18% of total sales, contribute 41.7% of revenue. Prioritizing this segment alone might limit future revenue due to a lack of ongoing sales opportunities.

Behavioral Segmentation:

Customers are categorized into two groups: Content-motivated customers who value learning over cost and cost-motivated customers. Content-motivated learners are more engaged and spend more.

Geographical Segmentation: Most customers are in the U.S./Canada (49.7%), followed by Europe (12.2%). This suggests potential for market expansion and targeted campaigns in specific regions.

Subscription Type Segmentation:

A notable disparity exists between limited (83.1%) and lifetime (17.9%) subscribers.

Customer Segments for Additional Products:

- A segment (9.7% of customers) opening and actively engaged with email marketing shows potential for additional product sales, but current email strategies are underperforming.

Strategies for Additional Sales:

- Implement targeted marketing based on customer location and language preferences.

Barriers to Engagement:

- Non-renewing customers show lower email engagement and are less likely to subscribe to emails with lowered engagement opportunities.
- To increase engagement, converting these customers to email subscribers will likely be the most effective.

Seasonal Trends:

- Increased activity in February and March presents opportunities for targeted marketing to boost subscriptions.

Recommendations:

- Balance efforts between converting existing customers to ALL-Lifetime subscriptions and acquiring new customers to replace converted customers.
- Utilize tailored marketing strategies based on customer behavior and geographical data to enhance conversion and retention.
- Leverage seasonal trends for strategic promotional campaigns.

REFERENCES

https://drive.google.com/drive/folders/1vk99yne3hJ4JKa2h_yQIlmhakW5-PWcn?usp=sharing

https://mailchimp.com/resources/email-marketing-benchmarks/

UNUSED VISUALIZATIONS

```
Auto Renew
Off     41.058814
On      18.087220
Name: Send Count,
```

```
Auto Renew
Off     39.749314
On      54.325813
Name: Purchase Amount,
```

```
Auto Renew
Off     54.943348
On      60.790260
Name: Purchase Amount,
```

```
Auto Renew  Email Subscriber
Off         No                  10954
            Yes                  8290
On          No                   6006
            Yes                  6816
```

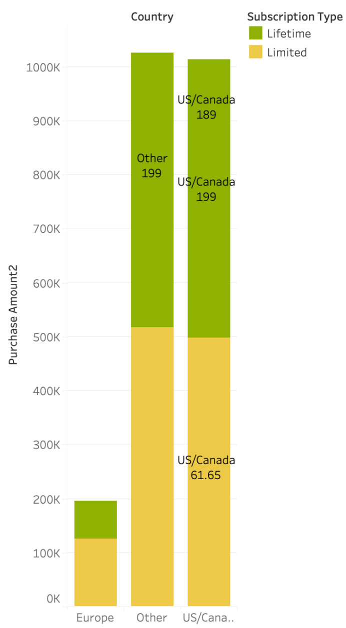
Data Summary for WEB only no cleaning.

	ID	Purchase Amount	Send Count	Open Count	Click Count	Unique Open Count	Unique Click Count
count	25294.000000	25294.000000	18504.000000	18504.000000	18504.00000	18504.000000	18504.000000
mean	22093.958607	69.157146	38.465413	10.534263	2.99265	4.764483	0.485409
std	11477.577636	77.339979	65.736794	44.819986	36.16901	14.482064	1.337054
min	2.000000	0.000000	1.000000	0.000000	0.00000	0.000000	0.000000
25%	12527.250000	0.000000	4.000000	0.000000	0.00000	0.000000	0.000000
50%	23300.000000	35.970000	11.000000	2.000000	0.00000	1.000000	0.000000
75%	32121.750000	119.000000	45.000000	7.000000	1.00000	3.000000	1.000000
max	39998.000000	996.030000	4370.000000	4365.000000	4348.00000	196.000000	44.000000

Averages of each language and type. We can use these values to fill in the 0 and or the App purchases.

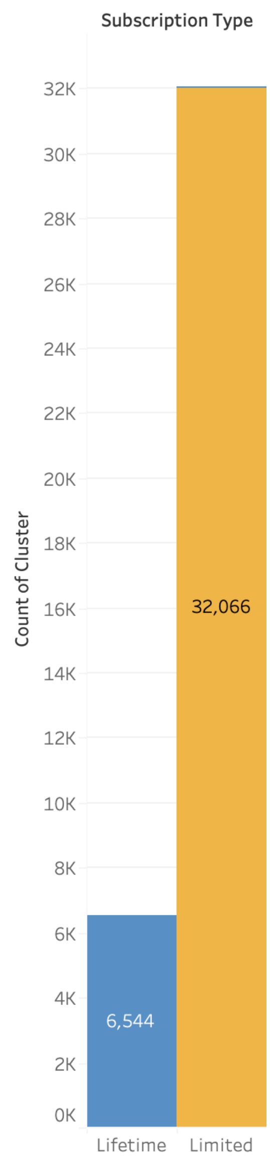
		ID	Purchase Amount	Send Count	Open Count	Click Count	Unique Open Count	Unique Click Count	Month	Year
Language	**Subscription Type**									
ALL	**Lifetime**	21704.888889	203.766667	53.826742	20.075565	5.669492	9.258945	1.079331	2.798542	2019.982549
	Limited	20976.878893	126.136021	29.204322	12.121807	5.161100	4.518664	0.612967	3.060554	2019.832180
ARA	**Lifetime**	20860.470588	166.817647	7.600000	5.600000	0.600000	1.733333	0.200000	6.294118	2019.529412
	Limited	22226.346154	46.833182	23.234694	4.311224	1.005102	1.887755	0.260204	6.339161	2019.017483
CHI	**Lifetime**	19465.681818	148.939091	5.315789	3.263158	0.947368	1.421053	0.263158	8.181818	2019.181818
	Limited	22596.729508	50.166448	22.496000	6.988000	2.372000	2.540000	0.340000	6.631148	2019.040984
DEU	**Lifetime**	17707.481481	165.438519	5.500000	2.545455	0.409091	1.136364	0.136364	9.037037	2019.111111
	Limited	20840.991696	50.350225	21.709677	7.654122	2.025090	2.885305	0.401434	6.190985	2019.091340
EBR	**Lifetime**	16969.727273	145.818182	9.266667	6.800000	4.666667	1.800000	0.800000	5.272727	2019.590909
	Limited	21608.893333	36.925552	24.269841	13.267196	7.960317	2.714286	0.531746	6.464242	2018.952727
ENG	**Lifetime**	24600.333333	149.206667	4.428571	3.000000	0.428571	0.857143	0.142857	8.333333	2019.111111
	Limited	20754.816794	53.717993	25.838352	7.920761	2.410460	2.832013	0.312203	6.499455	2019.107961
ESC	**Lifetime**	19035.478261	152.553478	8.500000	4.250000	1.300000	1.850000	0.400000	7.565217	2019.217391
	Limited	21791.234927	39.171320	21.003984	8.768924	2.800797	2.486056	0.380478	5.854470	2019.087318
ESP	**Lifetime**	22804.544715	175.127398	8.925532	4.680851	1.361702	1.819149	0.351064	6.682927	2019.390244
	Limited	21344.906667	61.344393	24.443975	5.728319	1.229087	2.688795	0.270913	6.386992	2019.183740
FAR	**Lifetime**	21437.000000	149.000000	35.000000	33.666667	20.333333	11.000000	1.666667	7.000000	2019.333333

Subrscition SUM Purchase amount and Purchase Type by Language



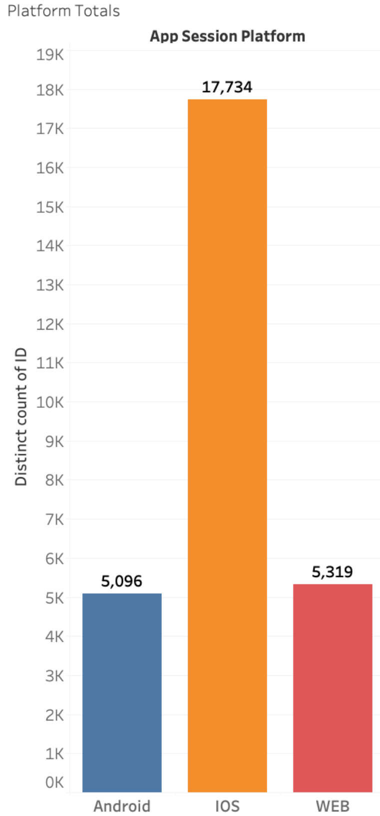
Sum of Purchase Amount2 for each Country. Color shows details about Subscription Type. The marks are labeled by Country and Purchase Amount.

Count of Users by Cluster and Subscription Type

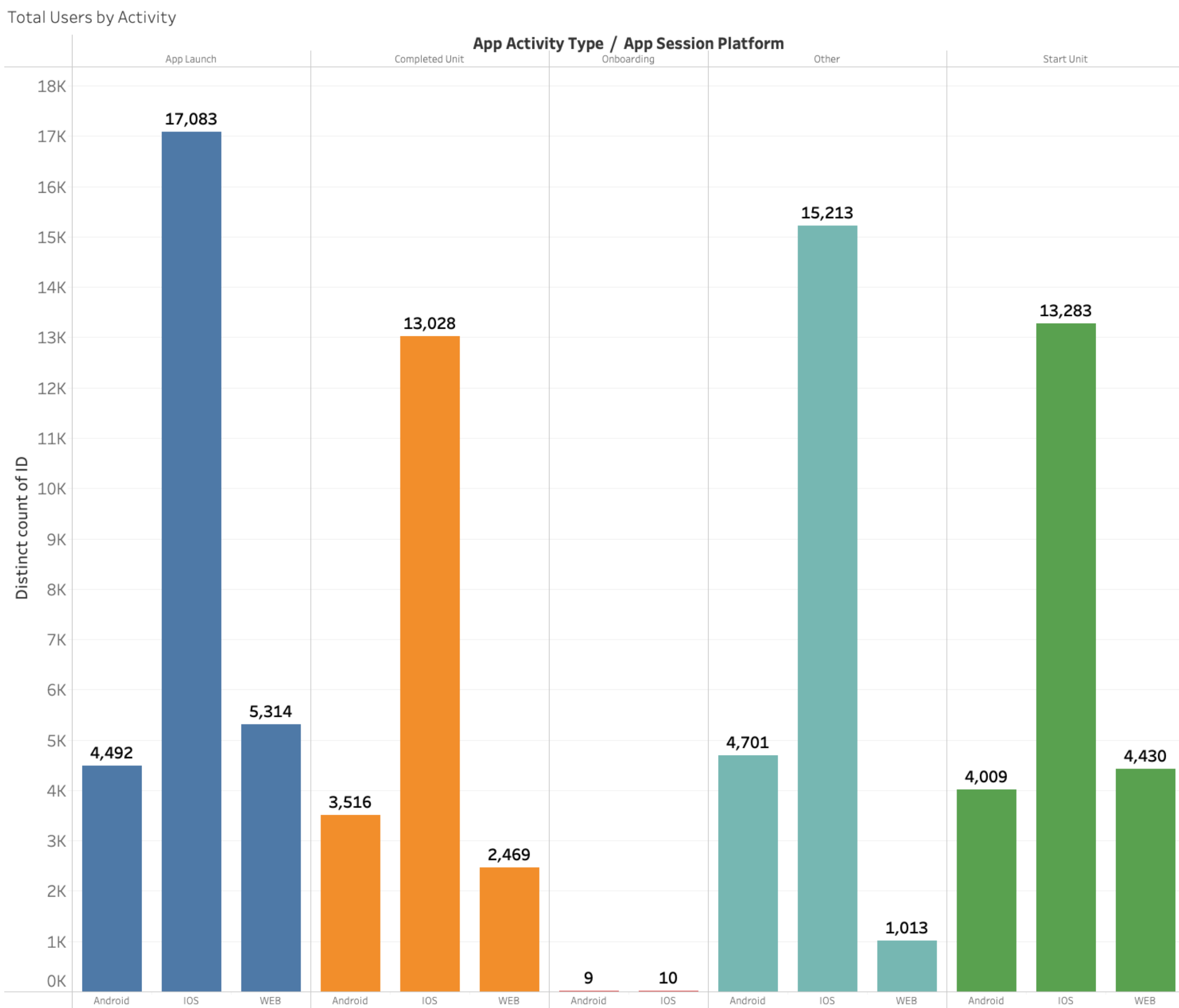


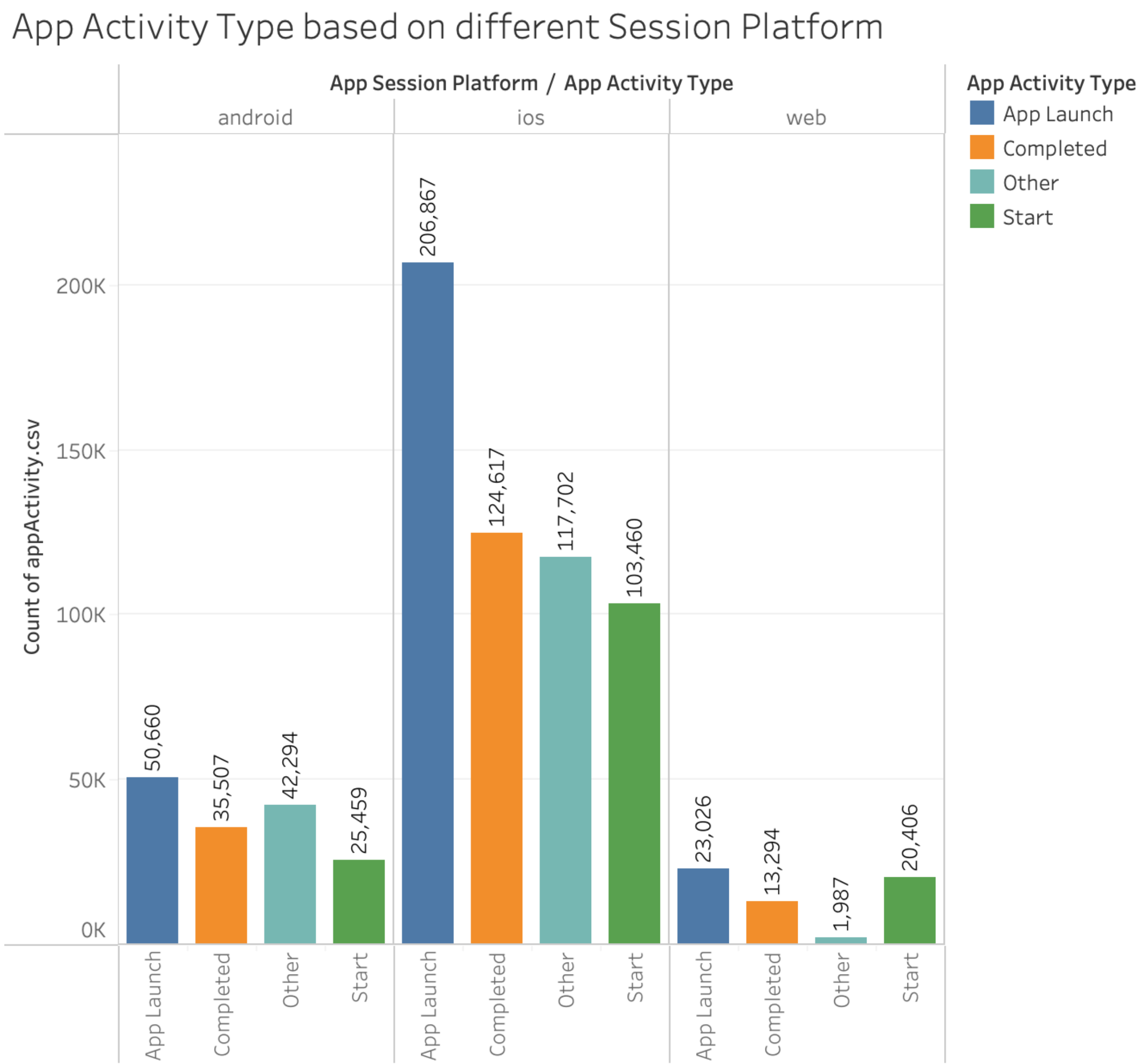
Count of Cluster for each Subscription Type. Color shows details about Cluster. The marks are labeled by count of Cluster.

Platform Totals



Data Assessment: App Activity

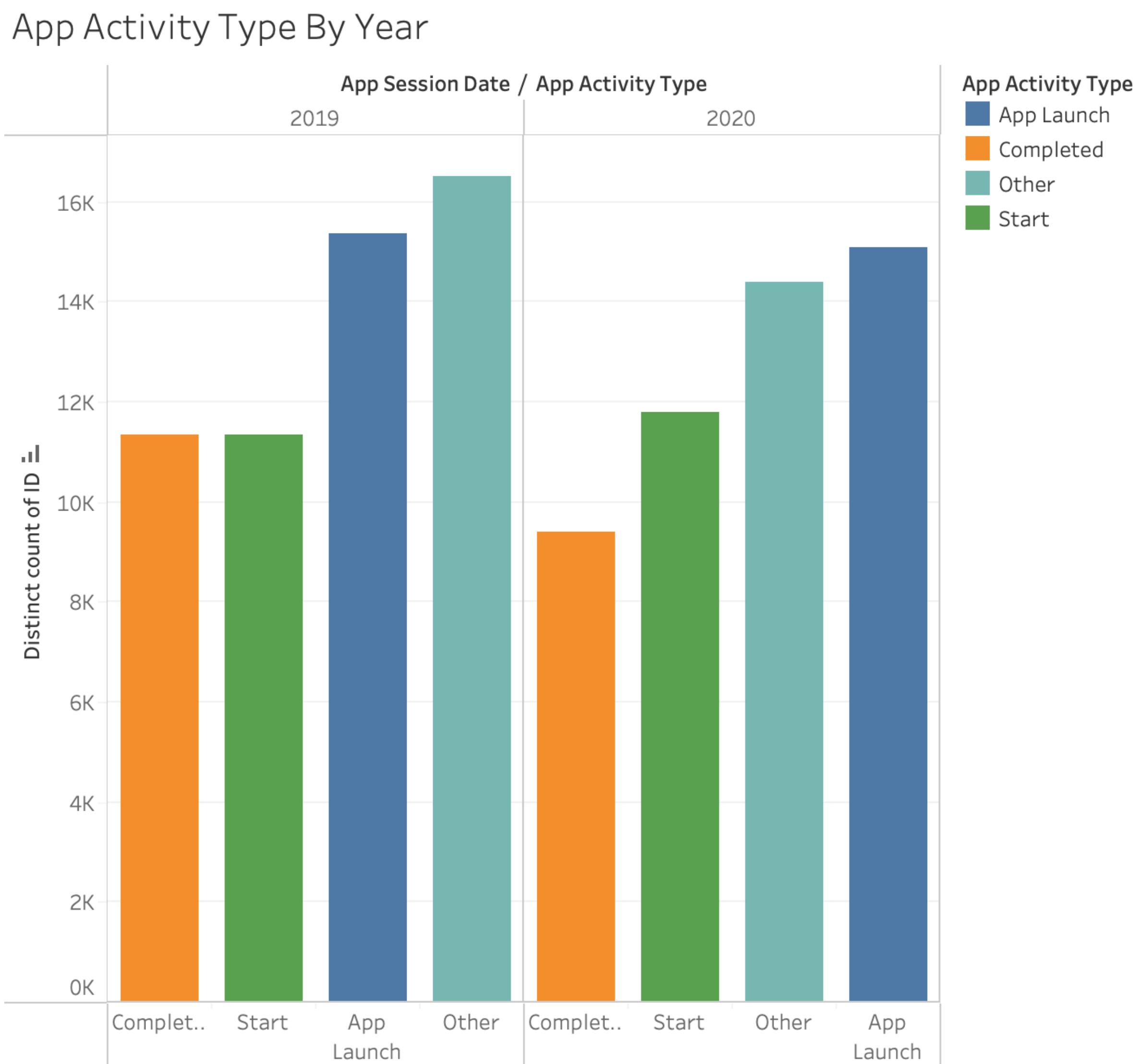

Total Users by Activity
App Activity Type / App Session Platform
App Launch
Completed Unit
Onboarding
Other
Start Unit
Distinct count of ID
18K
17K
16K
15K
14K
13K
12K
11K
10K
9K
8K
7K
6K
5K
4K
3K
2K
1K
0K
4,492
17,083
5,314
3,516
13,028
2,469
9
10
4,701
15,213
1,013
4,009
13,283
4,430
Android
IOS
WEB
Android
IOS
WEB
Android
IOS
Android
IOS
WEB
Android
IOS
WEB



Count of appActivity.csv for each App Activity Type broken down by App Session Platform. Color shows details about App Activity Type. The marks are labeled by count of appActivity.csv. The view is filtered on App Session Platform and App Activity Type. The App Session Platform filter keeps android, ios and web. The App Activity Type filter keeps Null, App Launch, Completed, Other and Start.

App Activity Types on different Session Platforms

This figure shows the App Activity Type based on different session platform.The platforms are android, ios, and web and the different activity types shown on the bottom are app launch, completed, other, and start. From this we can see that the IOS users were very high with app launch (opening the app), Complete (lessons), Other (other activities on the app), and Start (lessons). From the graph we can see that

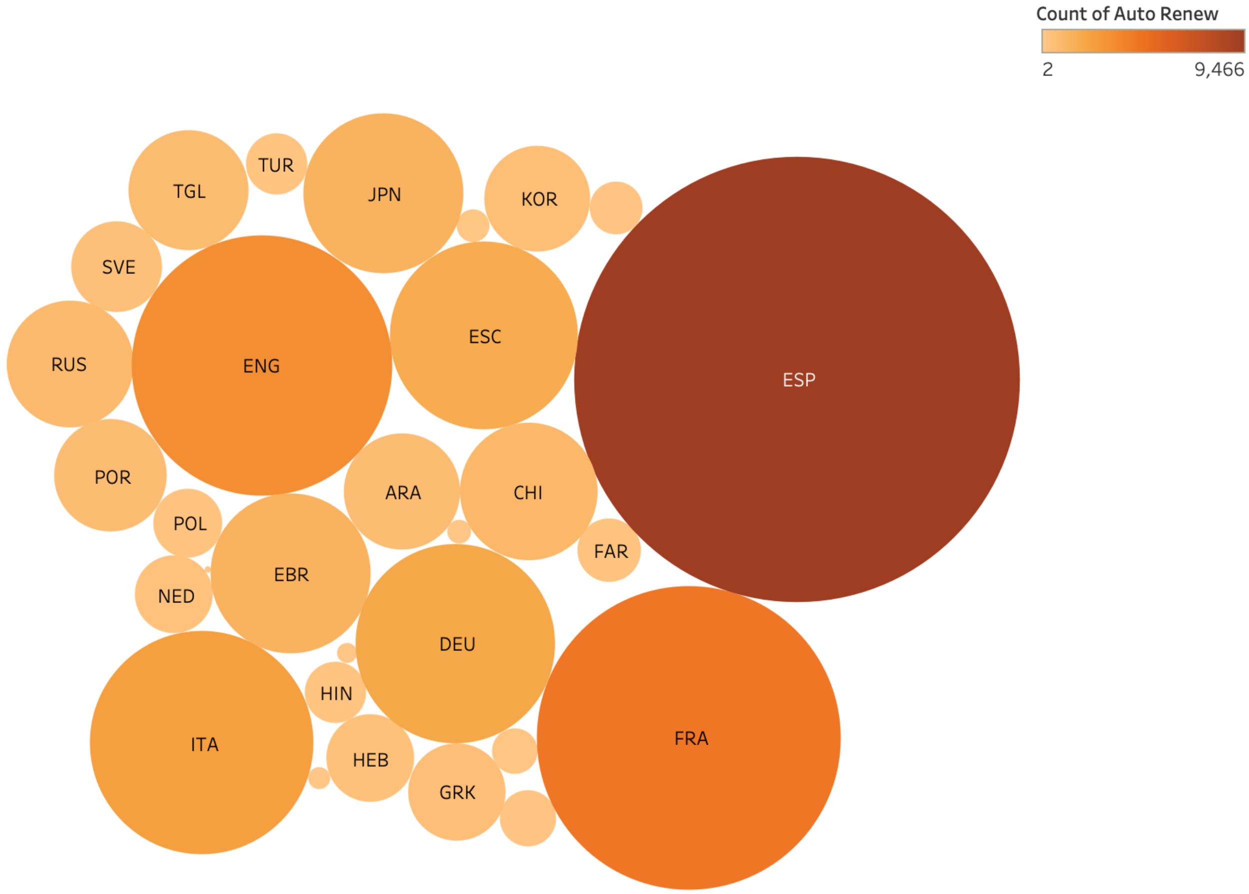


Distinct count of ID for each App Activity Type broken down by App Session Date Year. Color shows details about App Activity Type. The view is filtered on App Session Date Year and App Activity Type. The App Session Date Year filter keeps 2019 and 2020. The App Activity Type filter keeps Null, App Launch, Completed, Other and Start.

App Activity Types per year 2019-2020

This tableau visualizations shows the app activity type from a year to year standpoint. From this we can see there was more app launch activity in both 2019 and 202 than other app activities. In 2020 we can see that more people are starting lessons than completing lessons as well. However in 2019 completing and starting lessons looks very equal to one another.

Top Languages CNT Auto Renew

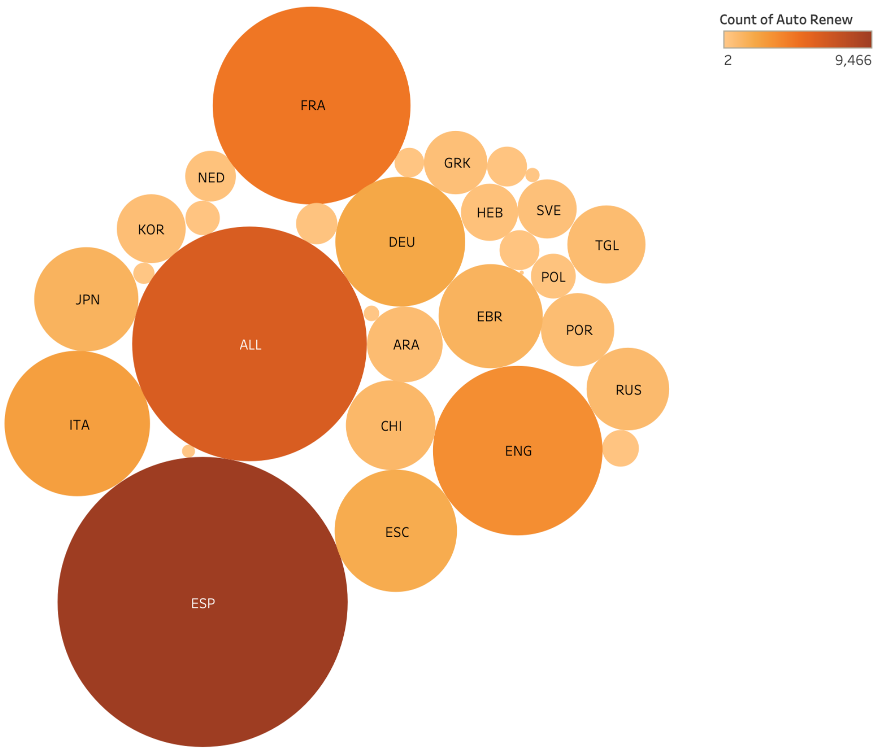


Language. Color shows count of Auto Renew. Size shows count of Auto Renew. The marks are labeled by Language.
The view is filtered on Language, which excludes ALL.

The visual is a bubble chart titled "Top Languages CNT Auto Renew," representing the count of auto-renewal subscriptions for various languages offered by Rosetta Stone. The size of each bubble corresponds to the count of auto-renewals for a specific language, with larger bubbles indicating a higher count. The color intensity represents the same metric, where darker shades suggest more auto-renewals.

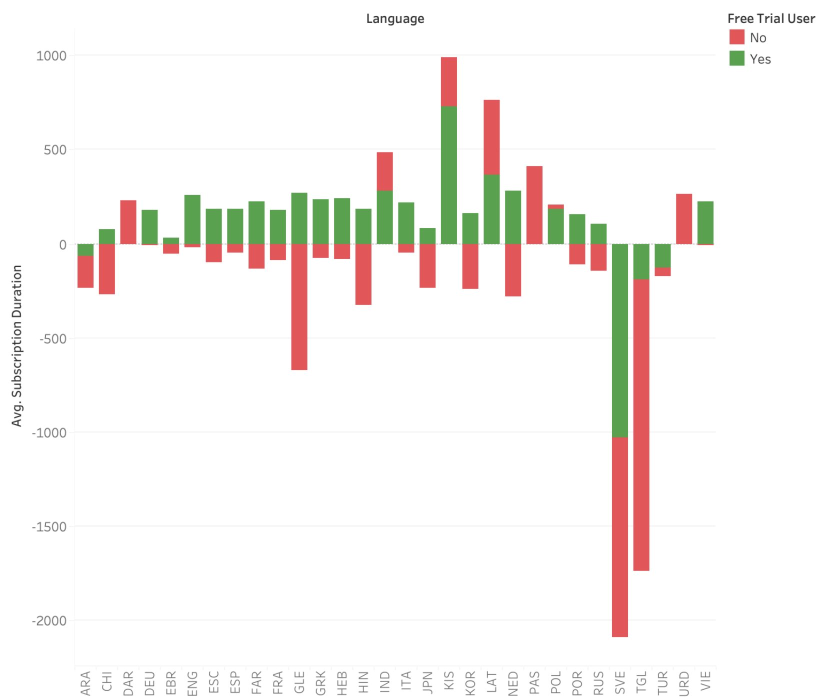
Spanish (ESP) has the largest bubble, indicating it has the highest count of auto-renewals among the languages, followed by French (FRA). Other languages represented include Tagalog (TGL), Turkish (TUR), Japanese (JPN), Korean (KOR), Swedish (SVE), Russian (RUS), Portuguese (POR), Polish (POL), Dutch (NED), Hebrew (HEB), Italian (ITA), German (DEU), Arabic (ARA), Chinese (CHI), Farsi (FAR), English (ENG), and Greek (GRK), among others. The count of auto-renewals ranges from 2 to 9,466. The chart excludes aggregate data, focusing instead on individual languages.

Top Languages CNT Auto Renew with 'ALL'



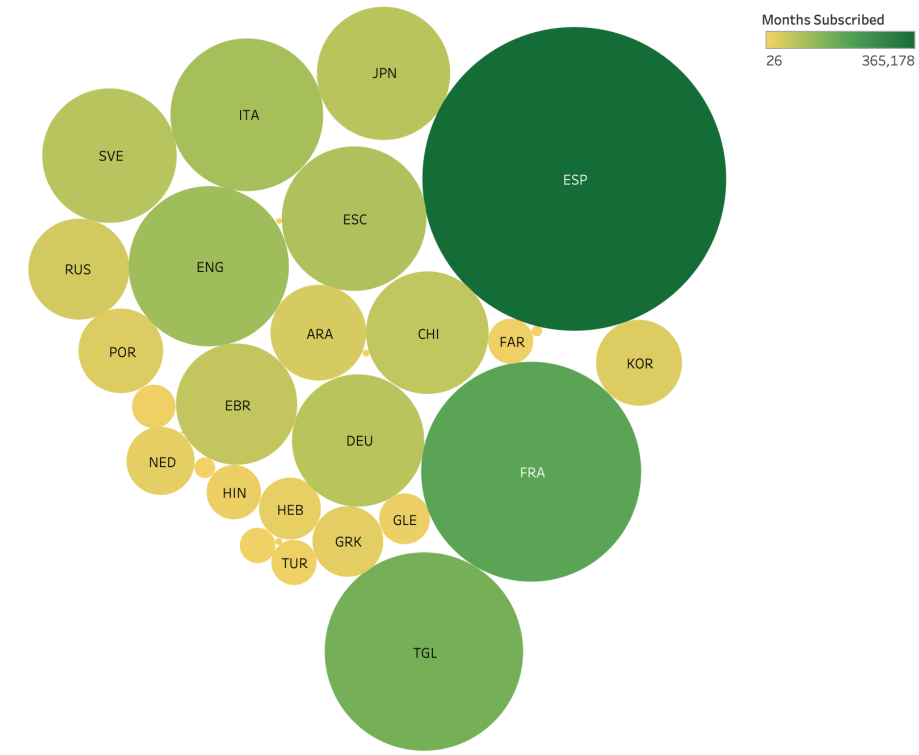
Language. Color shows count of Auto Renew. Size shows count of Auto Renew. The marks are labeled by Language. The view is filtered on Language, which keeps 31 of 31 members.

Average Subscription Duration by Language and IF Free Trial



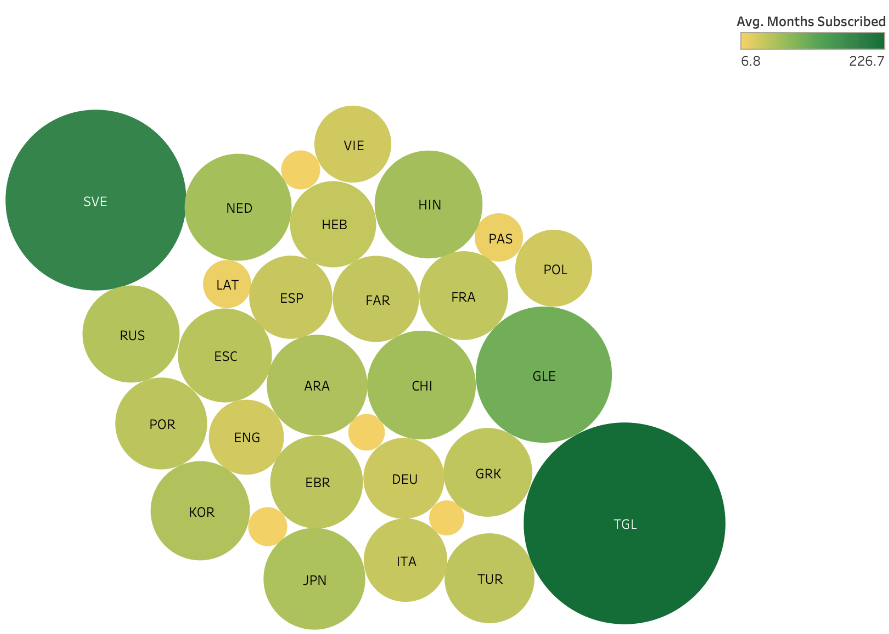
Average of Subscription Duration for each Language. Color shows details about Free Trial User. Details are shown for Language. The view is filtered on Language, which excludes ALL.

Top Languages SUM Months Subscribed

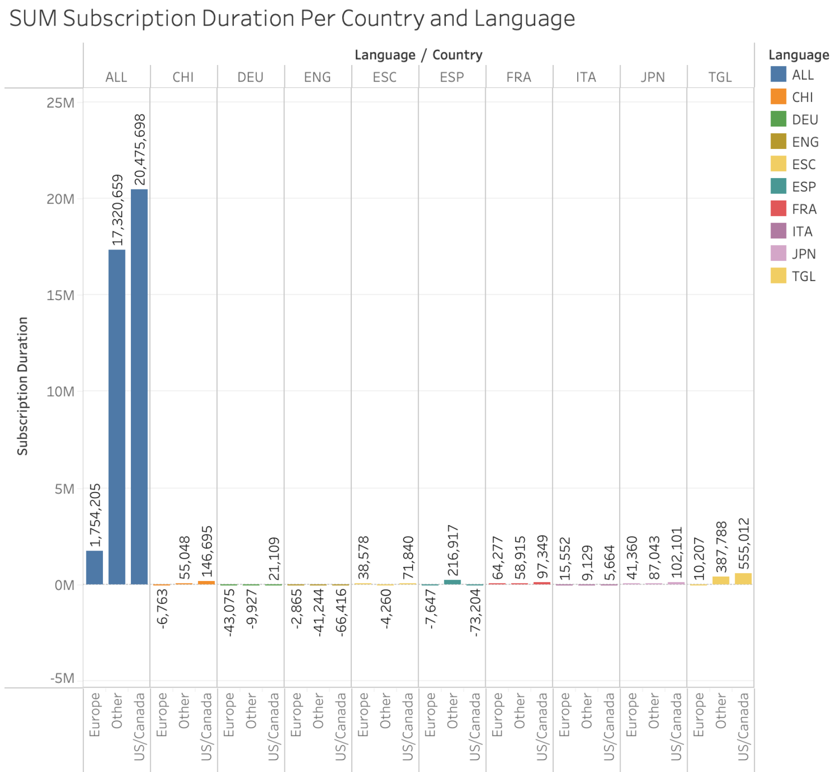


Language. Color shows sum of Months Subscribed. Size shows sum of Months Subscribed. The marks are labeled by Language. The view is filtered on Language, which excludes ALL.

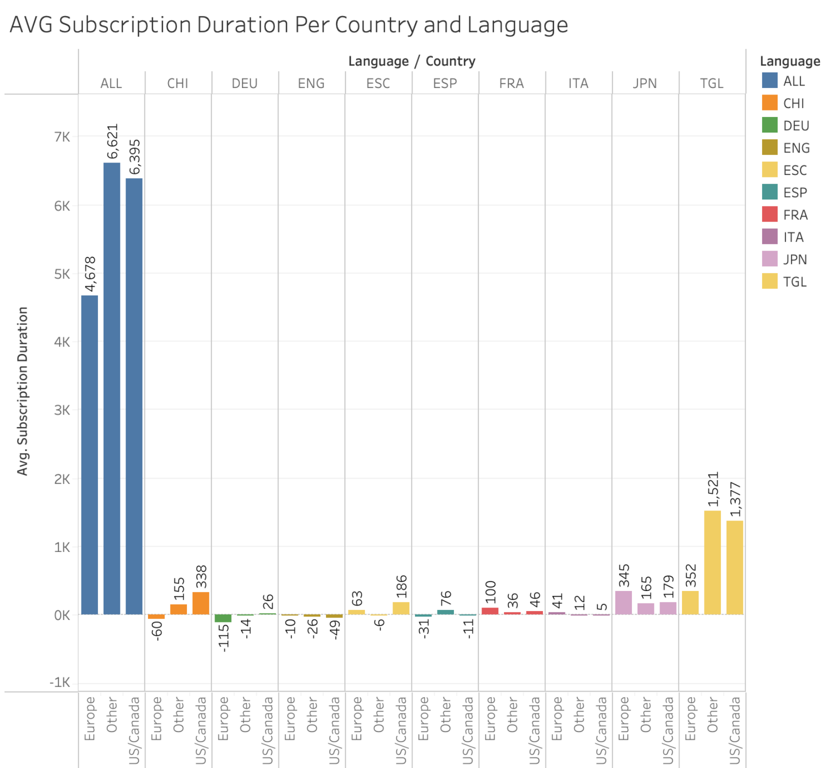
Top Languages AVG Months Subscribed



Language. Color shows average of Months Subscribed. Size shows average of Months Subscribed. The marks are labeled by Language. The view is filtered on Language, which excludes ALL.

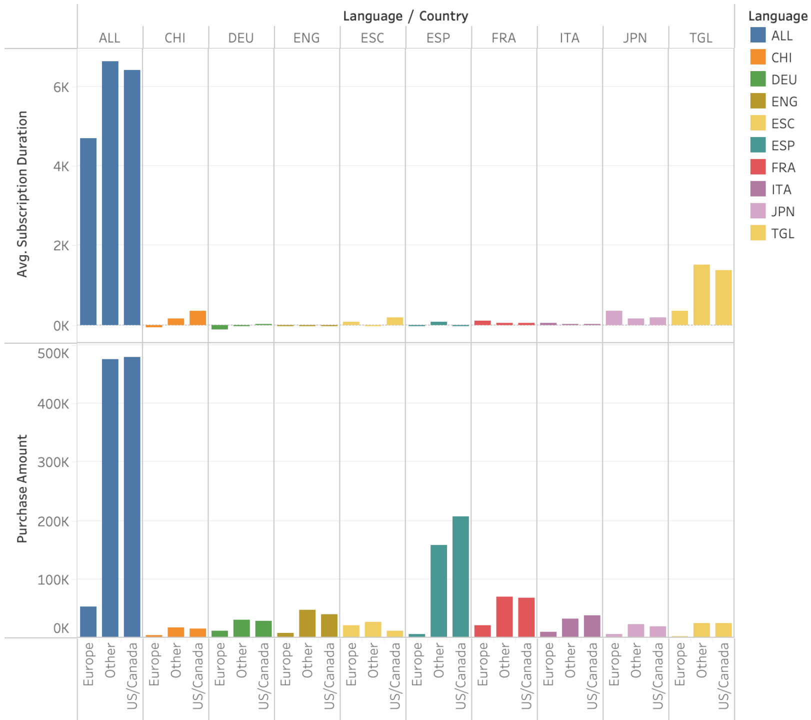


Sum of Subscription Duration for each Country broken down by Language. Color shows details about Language. The marks are labeled by sum of Subscription Duration. The view is filtered on Language, which keeps 10 of 31 members.



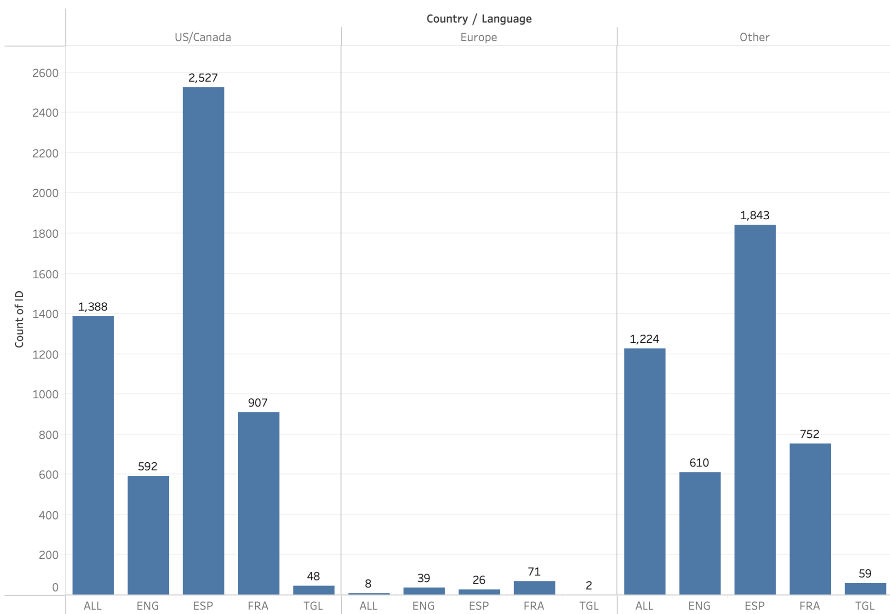
Average of Subscription Duration for each Country broken down by Language. Color shows details about Language. The marks are labeled by average of Subscription Duration. The view is filtered on Language, which keeps 10 of 31 members.

AVG Subscription Duration Per Country and Language with Purchase Amount



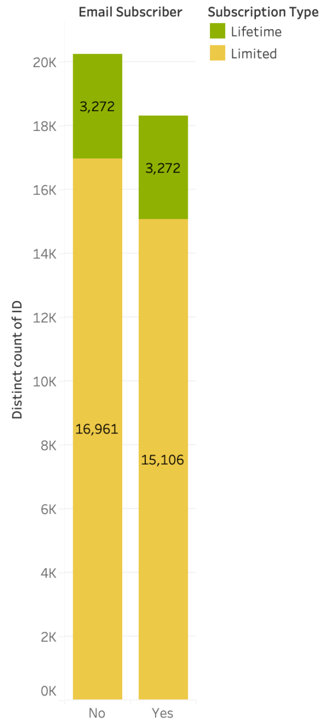
Average of Subscription Duration and sum of Purchase Amount for each Country broken down by Language. Color shows details about Language. The view is filtered on Language, which keeps 10 of 31 members.

Subscriber Amount Country/ID



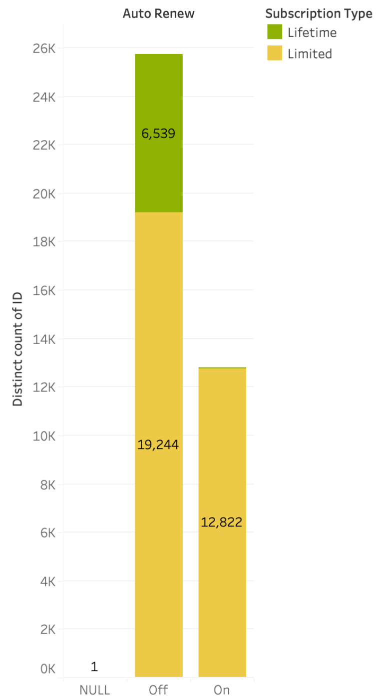
Count of ID for each Language broken down by Country. The marks are labeled by count of Purchase Amount (copy). The data is filtered on Purchase Amount (copy), which keeps 11 of 1,043 members. The view is filtered on Language, which keeps ALL, ENG, ESP, FRA and TGL.

Subscription Type; Email Subscriber



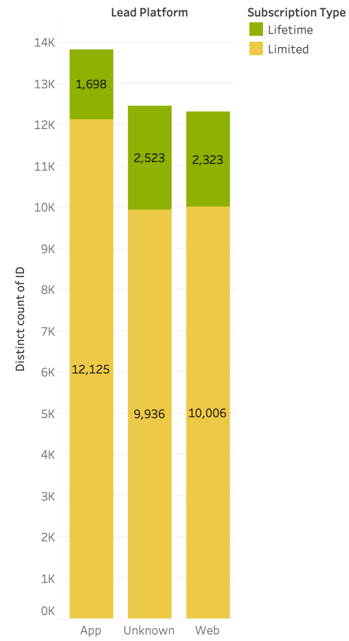
Distinct count of ID for each Email Subscriber. Color shows details about Subscription Type. The marks are labeled by count of Email Subscriber.

Subscription Type; Auto Renew



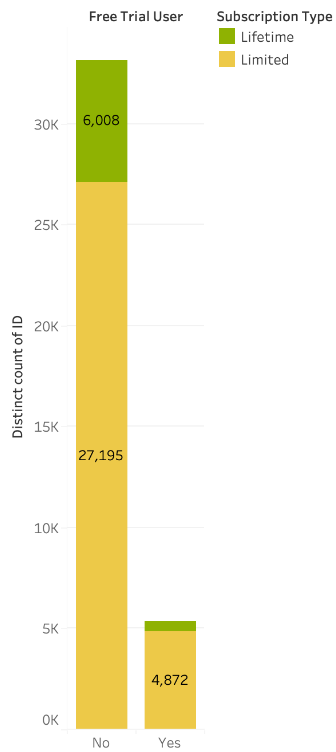
Distinct count of ID for each Auto Renew. Color shows details about Subscription Type. The marks are labeled by count of Auto Renew.

Subscription Type; Lead Platform



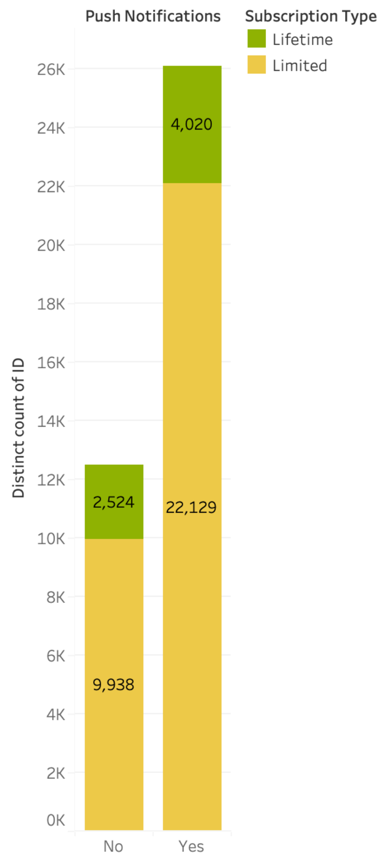
Distinct count of ID for each Lead Platform. Color shows details about Subscription Type. The marks are labeled by count of Lead Platform.

Subscription Type; Free Trial User



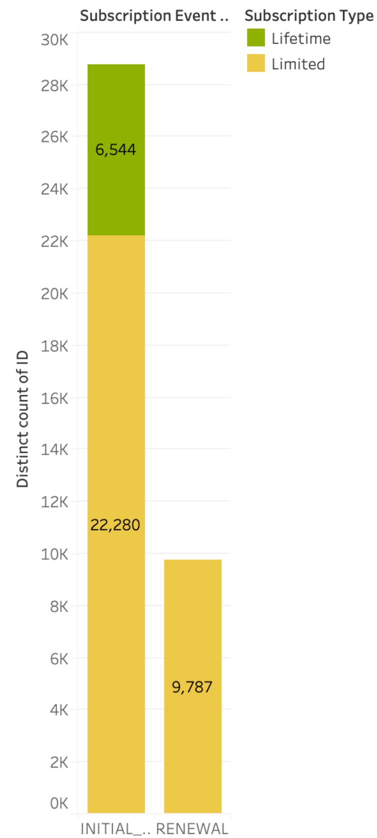
Distinct count of ID for each Free Trial User. Color shows details about Subscription Type. The marks are labeled by count of Free Trial User.

Subscription Type; Push Notifications



Distinct count of ID for each Push Notifications. Color shows details about Subscription Type. The marks are labeled by count of Push Notifications.

Subscription Type; Subscription Event Type



Distinct count of ID for each Subscription Event Type. Color shows details about Subscription Type. The marks are labeled by count of Subscription Event Type.