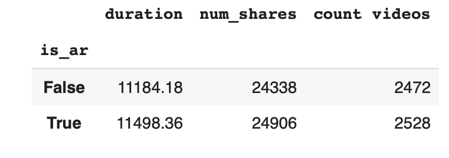
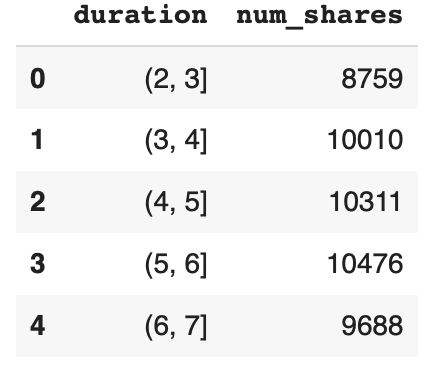
**Krikey Code Challenge**

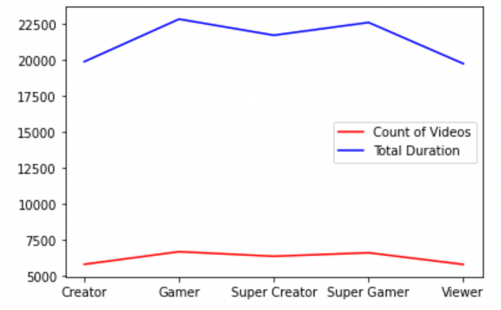
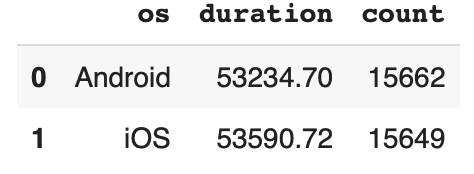
**Part 1: Exploratory Data Analysis**

We were given four datasets, which are users, videos, feeds, and watched videos. I mostly started the analysis by merging one or more datasets with the others to support deriving more useful information. On this part, what I mostly did is see the distribution between some columns, produced data visualization to see any trends, and derive conclusion based on that analysis. I added all the explanation on the code file below the analysis or visualization that also attached to that file. I started my analysis from basic questions, such as “How many users are there?” or “How many videos on a certain genre?”. To answer this question, I applied table manipulation and transformed it into visualization to make it easier to understand. After that, I deepened my analysis by diving more into the datasets and think what information could be useful for others.

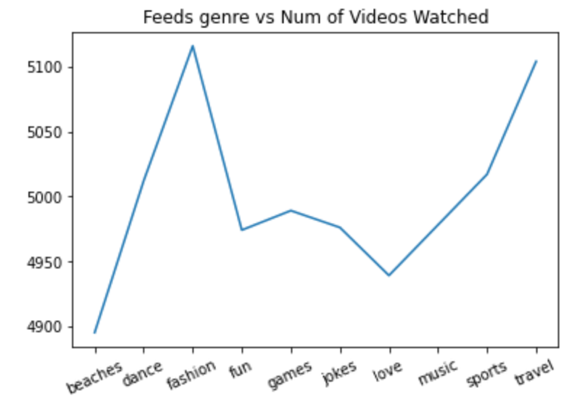
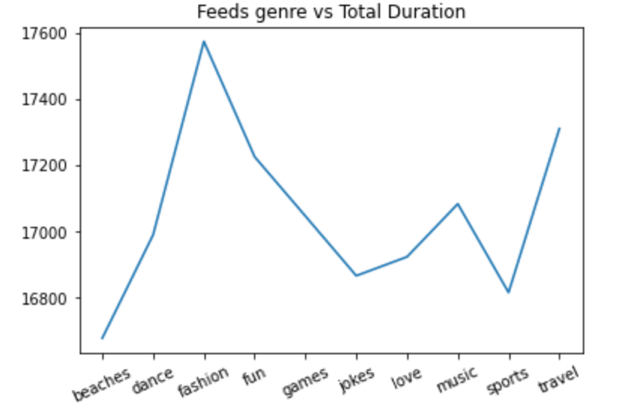
**Part 2: Two-Page Write Up**

Firstly, I explored the datasets using some Python built-in function such as info(), value\_counts(), and describe() to see any missing data, outliers, typos, etc. From here, I identified that all of the data has been cleaned. Thus, I started analyzing the datasets using Python (numpy and pandas) and matplotlib as data visualization. Here are my findings:

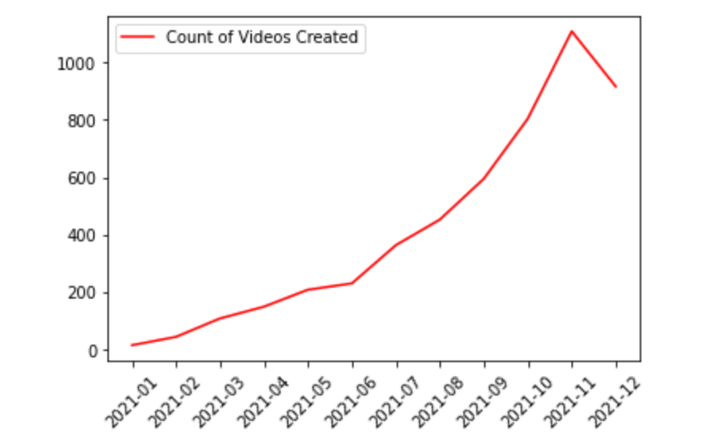
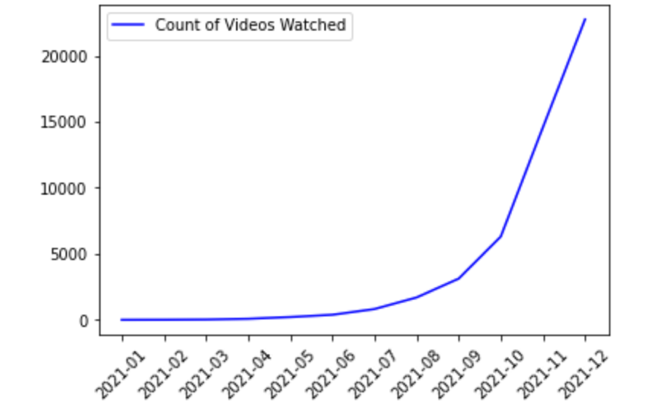
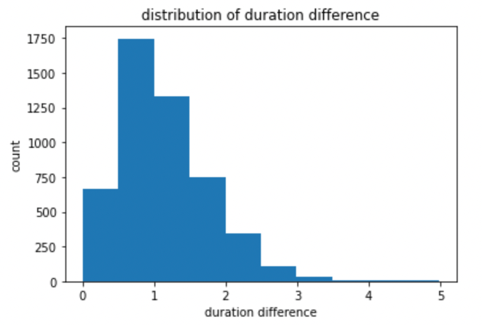
* From all the videos being uploaded, there are more videos being created using AR, and it leads to longer duration and higher number of sharing.
* The duration of the videos that would result on higher number of sharing is between 5 to 6 minutes. Thus, it is recommended to not create a video with too short or too long duration.
* Gamer watched the highest number of videos and so also spent the longest duration



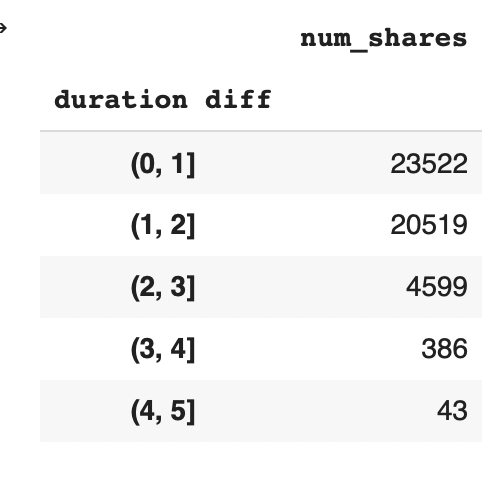
* The distribution of the watcher using Android and iOS are pretty similar from both the number of videos seen and the duration.
* The interesting part is that the most popular genre from the videos posted is fashion and travel, but the most watchers come from gamers, which logically would like to see gaming contents.
* Even though travel is one of the most popular genres, the duration of people watching the videos is similar to fun genre which has lower number of videos being watched. This means that we need to review the quality content of travel genre. On the other hand, once people watched fun videos, they seem to enjoy it resulted on longer duration of watching.



* The number of videos watched increased as the number of videos created also increased throughout the year.



* The distribution between the actual duration of the videos with the average duration of videos being watched being counted as duration difference. We can see that most of the watchers stopped watching the video around one minute before the video actually ended.



* Most of the watchers stopped watching video between 0-2 minutes before the video actually ended. This is a good number which means the video has a good enough quality content.

After performing data analysis and went through Krikey app, there are some changes that I believe would impact the product as a whole:

1. The videos dataset should also store the videos details, such as expressions type, dance, love, or sports, to see which animations are the most popular among the users. After that, we have an idea on how to update, advance, or change the animations.
2. Applying facial recognition tools to see the users’ demographic, such as gender and age. From that we can use this data to help us on decision-making on which target market we should focus on. This will impact where our advertisement budget will go into.
3. Store the number of likes of each video being posted to see what the trend is. We can connect it with the feed of the video to analyze what genre specifically is trending right now. Thus, we use this result to improve our app focusing on that genre.
4. Store the number of videos, followers, and following from each account and see if those number makes sense or it should be investigated further. This idea is to avoid hackers, scam, that would happen as the number of users increased.

The next steps for Krikey that should be explore more is expand their animation, features, and target market not only for India exclusively, but also globally. This is a new field that have a lot of potential, thus reaching out to kids and teenagers in terms of advertisement would definitely increase the number of users.

**Part 3: Krikey App Data Projects**

1. For each game (Run Ji Run, Yaatra Game, and Rangoli Puzzles), I would like to generate the number of users for each game throughout a year. If the number keeps increasing, we should focus on that certain game to be improved. On the other hand, for game that has a low number of interests, I would replace that game, so that the budget for that game can be allocated to the other area that have more room for improvement.
2. Analyzed the videos that are being reported, how many users report that video, and create a threshold to help decide whether the videos need to be taken down or not. The higher number of users, the more number of videos posted, the higher chance of reported video happened. To secure the company’s credibility, we want to satisfy our users as much as possible, and one way to do it is by taking down inappropriate video. Especially in this case, most of the users are kids, so we need to be careful on this area.
3. Created a summary of how many videos being uploaded, how many likes, and how many new users for each month and see how this number related to the company’s revenue. If there is a decrease at a certain time, we need to investigate why there is a significant decrease during that time. From this analysis, we can figure out where our revenue comes from and how to maximize it.