Jasmine Williams & Xavier Moore

04/23/21

Introduction to Data Science and Analysis Final Project

The Perfect Patreon



*“The Poor Poet” by Carl Spitzeg, 1839.*

Introduction

The above painting represents a problem that many artists and creators in the world have historically struggled with, and that is the act of living prosperously as an artist. However, in 2021 the Starving Artist can be a thing of the past, through the digital distribution service of Patreon. Patreon is a service that allows content creators to receive donations and profits in the form of tiered subscriptions in return for access to their work for the masses. However, not all Patreon accounts are created equally.

* What is Patreon?
* How do you use Patreon?
* Which categories are the most popular?
* Which subcategories are the most popular?
* How should you manage your content?
* How much money will you make?

Our project examines a dataset with the top 1,000 Patreons with categories such as art, music, podcasts, and videos, and numerous other subcategories. This is our unique project in which we will answer the above questions and go into further detail for our final project.

Prerequisites

* library(tidyverse)
* library(ggplot2)
* library(lubridate)
* library(stringr)

Why Did We Pick Patreon?

Upon meeting Xavier, I [Jasmine] was deciding if I was even going to submit anything for my final project and just receive a failing grade because I couldn’t find a group to work with. However, I met Xavier and he told me that he wanted to do his project on Patreon and explained to me why he felt that way. After listening to him, I decided that learning about Patreon wasn’t a bad idea and was actually something that I could learn from. I knew that people created podcasts, YouTube videos, etc. but never even heard of Patreon. Looking at the datasets that were already provided, I honestly began to second-guess my initial decision on whether or not doing a project on Patreon was a good idea. As you can see, we finished through with it and the finalized results will pay off in the end. Sticking with the datasets provided for Patreon was a bit of a struggle because they weren’t the easiest to deal with or analyze. Luckily, with our critical thinking skills, we were able to think outside the box in order to complete what we originally wanted to complete.

How Do You Define the Perfect Patreon?

The concept of perfection is a highly subjective and contestable thing. What does it mean to make a lucrative Patreon account? To answer this we came up with 51 different questions that can potentially be used to define this concept. Each question could only be answered after thinking critically in order to finalize the best questions for the project. We later narrowed down 51-question search to only 6 questions, in which are listed above. We made sure that our questions would be able to introduce the idea of Patreon to our audience, as well as providing compelling information. All in all, we were able to come to the conclusion of what type of content produces the most profits, what is the most popular, and how to manage your content.

Dataset

The first and arguably most important piece of information that we sought to answer for this project was “What category of content was most popular?” Logically, chasing trends while they’re trendy increases the odds of you succeeding quickly. We wanted to answer that question but upon loading and evaluating the .csv file we saw that it was for, lack of better term: “Jacked-up!”. This was because the spreadsheet recorded most of the information as character values, including things that should have been numeric and was missing some stuff that would have made understanding the content easier like Genre Category Labels. So, the actual first thing we did was try to “Un-jack-up” the spreadsheet. To that end we looked through the content and began discovering a number of keywords and phrases associated with one of 4 main and common content categories for every Patreon Account. The categories were: art, music, podcasts, and videos. To assign concise categorizing variables, we defined several character arrays for each genre containing the aforementioned keywords, we then filtered the keywords in the arrays using an *‘str\_detect’* function to identify each row with the corresponding content and then used the *‘mutate’* to and *‘bind’* function to create a new column properly labeling all the rows by content. After that, it was a simple matter of *‘filter*’, *‘group\_by’*, and *‘summarise’* being employed to identify the most common and popular Category of Content overall. Then you just had to define a new dataset with the most popular content filtered and do the steps all over again post “Un-jacking-up” only this with the intent to find the best subcategory of content. The results of our research was the best thing to make a Patreon account for Video Game, Videos, on YouTube or a Podcast on YouTube.

The second step was to analyze the trends when it came to money and profits. The columns which displayed this information were unfortunately also “Jacked-up”, so to make the information usable, we had to use the old *‘mutate’* function once more, this time with the *‘str\_extract’* function and special parameters to pull the digits out of the character string and store them as purely numerical values. We performed this process on the “Per Patron” and “Earning” Column respectively. Then it was a simple manner of once again using the *‘summarise’* function to find the averages and then *‘ggplot()’* to visualize potential relationships and correlations between these values.

Conclusion

As aforementioned, Patreon is a service that gives artists an opportunity to make a living as they create. We realized that the idea of being able to make a living for yourself as an artist/independent creator in this fashion is truly fascinating, especially since a number of services exist now to start-up your identity on platforms like YouTube, SoundCloud, and DeviantAr in the modern world. The questions we chose were what we believed would answer the main postulation of “What is the best way our audience could make a living off of Patreon if they so choose?” and present the ideas of Patreon in what is hopefully the most informative and interesting ways to our audience. This project displays the trends that point to the best way to profit off of Patreon’s services: to create content for the audience, as well as providing a monthly income for the creators.

As for what we discovered during this project, the most popular categories from the most to least popular are: videos, podcasts, art, and music. Likewise subcategories, it is YouTube videos, video games, ASMR. What we found to be interesting is that just from making content from the categories listed, you can easily make $13,000+ annually just by being consistent and producing the type of content that is wanted by the subscribers. So, to best succeed at patreon you should be a YouTuber with a channel dedicated to Video Games, Podcasts, or both. The average you should charge for your Patreon after you are established is $4.25 across your tiers and marginally increase this amount with the number of patrons you obtain. Update your account with content regularly and consistently over 3-4 years and you will be rolling in dough in no time.

What Are We Going to do With This Information?

After learning what we did from this project, we found it fascinating how we could reasonably make a profit from becoming YouTubers with Podcasts or Let’s Play Channels if we really wanted to. If we were to do this, there is a proportion between the content presented and the amount of money that can be charged. For example, I mentioned if we were to make a YouTube page, we can’t charge $15 for a subscription and only post two videos. However, what we can do is require a reasonable subscription such as $4-5 per month. This is a price that tends to be low enough to maintain subscribers. Additionally, we learned how to analyze data with predominantly, random character based objects and convert them into numerical and categorical information. Moreover, this same concept doesn’t only have to be applied towards similar platforms such as YouTube and anything that profits from user based subscriptions..