90’s Disney Hits

Ryan Batten

10/05/2021

# Disney Hits + Spotify: All Together

The aim of this project started as a NCAA March Madness style tournament bracket. What began as a simple selection process, turned into a more data-driven goal of “this is good but how do we look at this with a scientist’s eye?”. Data are everywhere, but sometimes it requires a bit more of “playing around” to understand what is really going on here.

## Step 1: Getting Setup with Spotify

If you have no clue how to do this, don’t worry! I didn’t either until I started playing around with it. First step is to head over to the amazing Thompson et al. (2021) *spotifyr* package to get setup with a Spotify developer account.

## Step 2: Getting Familiar with the Data

So now you’re in, what next? When you start exploring with the data it can seem overwhelming. Lists on lists on lists….but not to fear young Jedi! Where there is a will there is a way. This is the exploratory data analysis phase, more of an art-form than a science however can’t recommend enough to check out “R for Data Science” by Grolemund & Wickham (2016). Like many before me, this is a great place to start as a beginner.

We’ll pick the four albums we want to examine (using Disney Hits we all know and love!): The Lion King, Hercules, Tarzan and The Little Mermaid.

## Step 3: Alright, now you’re familiar what’s next?

This part requires a little more googling (a key part of everyone’s process: from data scientists to top level execs) to understand what is exactly going on here. Once you have a Spotify developer account, feel free to start poking around about what each component means: how the data is stored, how to access it, etc. Spotify has some great, very detailed guides on every aspect of what is in store.

## Step 4: Lets Get the Album ID

If you have the Spotify app already downloaded onto your desktop (like I do), click on the album you’d like to use to pull the data from. You will see three dots (…): click share then “Copy Album Link”. This will give you the unique Spotify album ID. Again, using The Lion King (1994) we end up with: “<https://open.spotify.com/album/3YA5DdB3wSz4pdfEXoMyRd?si=CpCUiyLBT2ea5y7haYA9Cg>”. The last string, that bizzare set of alphanumeric characters, after the “album/” are what is referred to as the Spotify album id.

## Step 5: Data Collection

Time to get into it young grasshopper! We’re going to use the *spotifyr::get\_album()* function to extract what we need here. It will pull a **BUNCH** of meta-data but fear not. If we keep this simple, we’ll be just fine. Below is an example using The Lion King (1994)

lion\_king <- spotifyr::get\_album(id = "3YA5DdB3wSz4pdfEXoMyRd?si=9UUAysDuTS6ixCsXYMykWg")

As you can see, each album is associated with a unique identifier (typically a string that consists of a bunch of nonsensical letters!). If they didn’t make sense to you, don’t worry they don’t to me either!

Here is a list of the other three albums:

tarzan <- spotifyr::get\_album(id = '6fBzYwBKjuO4hmhcGuklJM?si=8G4ZAzUVRcm0FRF8aXphuw') # Tarzan (1994)  
little\_mermaid <- spotifyr::get\_album(id = "4YTduhQWfS0pOzQC4o0HcG?si=LBpETiG2RSyZyIk8mloBQg" ) # The Little Mermaid (1997)  
hercules <- spotifyr::get\_album(id = "1wbY6VUchNsZLaDi22eD3J?si=00ajCSlHReuJXSz8XpARZg") # Hercules (1997)

## Step 6: Popularity

Now, we want to see how popular each of these songs are. There is a variable labelled: popularity, which ranges from 0 to 100 with 100 being the highest / most popular. Popularity is not static rather it is dynamic, so it is liable to change over time. Anyways, at the time of this writing: May 9, 2021 the results are (drummmmrollll pleaseee):

big\_finale <- dplyr::tribble(  
 ~album\_name, ~popularity,  
 "Tarzan (1999)", 1,  
 "The Little Mermaid (1997)", 65,  
 "Hercules (1997)", 67,  
 "The Lion King (1994)", 76  
)  
  
big\_finale <- big\_finale %>%   
 dplyr::arrange(  
 -popularity # organizing table by popularity  
 )

big\_finale\_table <- flextable::flextable(big\_finale) %>%   
 set\_header\_labels(album\_name = "Album Name",   
 popularity = "Popularity") %>%   
 flextable::theme\_zebra()  
  
big\_finale\_table

| **Album Name** | **Popularity** |
| --- | --- |
| The Lion King (1994) | 76 |
| Hercules (1997) | 67 |
| The Little Mermaid (1997) | 65 |
| Tarzan (1999) | 1 |