

PONTIFÍCIA UNIVERSIDADE CATÓLICA
DO RIO DE JANEIRO



Departamento de Economia

Monografia de final de curso

**O Conhecimento Sobre Os Artistas Continua Sendo O
Principal Fator Para Seu Sucesso No Meio de
Streaming**

Gabriel Konkier, 1612532

Orientador: Leonardo Rezende

02/07/2021

Declaro que o presente trabalho é de minha autoria e que não recorri, para realizá-lo, a nenhuma forma de ajuda externa, exceto quando autorizado pelo professor tutor.

PONTIFÍCIA UNIVERSIDADE CATÓLICA
DO RIO DE JANEIRO



Departamento de Economia

Monografia de final de curso

**O Conhecimento Sobre Os Artistas Continua Sendo O
Principal Fator Para Seu Sucesso No Meio de
Streaming**

Gabriel Konkier, 1612532

Orientador: Leonardo Rezende

02/07/2021

Declaro que o presente trabalho é de minha autoria e que não recorri, para realizá-lo, a nenhuma forma de ajuda externa, exceto quando autorizado pelo professor tutor.

As opiniões expressas neste trabalho são de opinião única e exclusiva do autor.

Agradecimentos

Primeiro agradeço a Puc-Rio, faculdade em que curso que me permitiu o aprendizado e disponibilizou seu corpo docente ao longo da minha formação para o ensino. Agradeço em conjunto a todos professores que estiverem presente ao longo da minha formação.

Gostaria também de agradecer minha família por sempre me ensinar e me guiar pelo caminho correto e apoiar nas minhas decisões ao longo da vida e do meu estudo. Obrigado ao meu pai, mãe, irmão e avós.

Agradeço a todos meus amigos que estiverem presente neste caminho, em especial Tamir, Victor Hugo e Bernardo Fernandes, sem eles não sei o que seria da minha formação e dos meus dias dentro da faculdade. Obrigado por todo o suporte e ensino quando foi necessário.

Obrigado ao meu orientador por estar disponível e disposto a abordar um tema que se afasta dos temas mais ordinários, me apoiou e guiou no caminho de elaboração do presente trabalho.

Por fim, obrigado a todos que ouviram, discutiram, questionaram ou participaram de qualquer forma desta monografia.

Sumário

1. Introdução	7
2. Revisão de Literatura	9
A. Música e a Internet	9
B. Streaming	10
C. Information and the Skewness of music Sales.....	11
D. Backward Spillover	12
E. Youtube	13
F. Tiktok	13
3. Metodologia e Dados	14
A. Seleção de Dados.	16
B. Problemas e Soluções	16
C. Quebras Estruturais.....	18
4. Análise	20
5. Conclusão	24
6. Referências Bibliográficas	25
7. Apêndice	27

Lista de Figuras

Figura 1: Receita gerada por músicas gravadas	11
Figura 1: Cannibal – Kesha Tiktok vs Youtube	15
Figura 1: Cannibal – Kesha Tiktok vs Youtube [correção I]	17
Figura 1: Cannibal – Kesha Youtube	18
Figura 1: Coño – Puri Youtube	19

Lista de Tabelas

Tabela 1: Análise numérica	20
Tabela 1 Análise numérica II	20

1. Introdução

Ao longo das últimas duas décadas viu-se o mercado musical mudar do meio de mídias físicas para o meio de mídias digitais. O método de *Streaming* de música se tornou o meio mais comum para o consumo de faixas de áudio. Antes disso, o mercado musical era composto principalmente pelo Rádio, CD's e vinis, porém o consumo de vinis foi substituído pelo de CD, o qual foi substituído pelo de streaming.

Para uma música se tornar popular em meios que não existiam a internet, ela necessitava tocar na rádio ou estrelar em paradas, o que não era comum para artistas completamente desconhecidos, uma vez que para alcançarem as paradas necessitavam já ser uma das músicas mais tocadas, portanto o principal meio para um artista “sair do zero” era através das transmissões de rádios populares. Isto criava uma espécie de *Bottleneck* informacional, já que não era possível colocar no ar todos artistas por falta de tempo. Com isso, muitos artistas nunca veriam a luz do dia pela indisponibilidade de espaço para o conhecimento e compradores eram relutantes em comprar álbuns de artistas desconhecidos, uma vez que tinha um custo.

No momento atual com o streaming e a internet, toda informação está disponível, o custo de ouvir um álbum ou infinitos álbuns é o mesmo, portanto poderia se esperar que artistas não encontrassem esse *bottleneck*. Porém a quantidade de artistas subiu, assim como a quantidade de música disponível, mais uma vez inviabilizando que todos artistas se tornem estrelas.

Neste estudo busco provar que uma externalidade pode ser muito mais eficiente em tornar um artista famoso do que a própria qualidade da sua música. Através da plataforma Tiktok e Youtube, observaremos que a música fazer sucesso no Tiktok, um aplicativo que depende de músicas de fundo para suas *trends*¹ porém limita a apenas 1 minutos da música, fazendo com que ele não seja um aplicativo de streaming de música, afeta na quantidade de visualizações do Youtube, um aplicativo que é usado para o streaming de música.

¹ Moda criada no Tiktok

Chamaremos este acontecimento de *Spillover*, que nada mais é do que uma externalidade que parece não relacionada. Ao encontrarmos isto temos como interesse argumentar em prol de uma afirmação feita em Goeree (2005): o rápido avanço da tecnologia leva a consumidores menos informados. Que mesmo com os meios de informações mais desenvolvidos e toda informação disponível a todos em todo momento, continuaremos deixando de ouvir artistas por não sabermos que eles existem.

Porém não é objetivo do trabalho discutir a fundo essa questão de forma quantitativa, apenas mostrar que os consumidores continuam não consumindo músicas por falta de conhecimento delas. Busca-se reencontrar os resultados de Hendricks e Sorensen (2009), porém desta vez atualizados para o *streaming* e tecnologias atuais.

Em Salganik, Dodds e Watts (2006) encontra-se resultados que apoiam nossas buscas. Ao criar uma plataforma digital para músicas onde os participantes são apresentados uma lista de músicas de artistas desconhecidos, cada usuário dará uma classificação para música e se queriam ouvir e baixar a música (sem custos), eles podiam ser informados ou não das escolhas dos outros participantes. Aqueles que foram informados das popularidades costumam ouvir apenas as músicas mais populares, porém invariante para baixar. Portanto eles escutavam quando os outros escutavam, porém baixavam se gostassem

“Medium-quality songs had the most unpredictable download totals: “the best songs never did badly and the worst songs never did well but any outcome was possible for songs in between.” (Hendricks, Sorensen 2009)

Na indústria musical, muitas músicas novas são apresentadas todos os dias e com isso o consumidor não é capaz de ouvir todas disponíveis, portanto as demandas não dependeram apenas do consumidor, mas sim também do espaço de mercado que o produto encontrar para se tornar conhecido.

2. Revisão de Literatura

A. Música e a Internet

Em pleno monopólio do CD, em 1999 aparece o Napster, o primeiro meio popular de baixar músicas online, criando uma nova forma de consumir música. Apenas quatro anos depois, em 2003, surge o iTunes, e os usuários se mostraram contentes em pagar por música, mostrando que a agora buscam apenas o imediatismo. Embora a pirataria seja um fator relevante, ela se mostra contornável, demonstrando mais uma vez que o usuário está mais preocupado com o imediatismo do que com o baixo valor que o meio online provê.

Com o sucesso da música via internet, novos serviços aparecem com propostas diferentes, como por exemplo: Last.fm, que utilizava algoritmos para oferecer músicas relacionadas com as que o usuário já consumia, Pandora, oferecendo estações de rádio personalizadas para cada usuário. Estes foram evoluindo até se tornarem o meio mais comum de consumo musical.

O youtube ganhou espaço no mundo como um serviço de vídeos, porém muito rapidamente artistas perceberam o quão importante era ter suas músicas disponíveis em uma das maiores plataformas digitais do mundo. Assim, até hoje, um dos principais meios de streaming de música.

Em um primeiro momento o iTunes, seus concorrentes legais e a pirataria se tornaram cada vez mais comuns, porém na última década perderam espaço para o meio de *streaming*, que consegue se rentabilizar através de assinaturas mensais ou propagandas. O Spotify, criado em 2008, tem um papel fundamental neste novo mundo, visto que, com mais de 170 milhões de usuários, se tornou o principal meio de *streaming* musical. Hoje, artistas não precisam mais investir em mídias físicas para a venda de seus fonogramas, e principalmente, não dependem mais de um álbum completo ou uma gravadora para lançar suas músicas.

A nova facilidade para lançamentos permitiu que artistas não precisem mais compor um álbum completo e tenta-lo *pitch*² a gravadoras. Não se faz mais necessário lançar muitas

² Promoção da música para rádios, *playlists* ou outros artistas

músicas de uma vez, há muitos artistas que lançam apenas *singles* e assim, *single hit artists*³ se tornam cada vez mais comuns.

B. Streaming

Também no início da década começaram os primórdios do *streaming*, mas não com muita popularidade. Nasce o Last.fm e Pandora, porém eram dependentes, assim como o rádio, da escolha musical de DJs. Apenas com Myspace e Souncloud que nasce o *streaming* moderno, onde cada um escolhe sua própria música no momento que quiser. Porém, ainda estava longe de ser a forma mais conveniente de se escutar música em movimento, pois a internet móvel não permitia isso.

A transição do meio físico para o meio digital da música inicia uma redução na receita das gravadoras e de músicas, ao invés de se comprar um álbum físico podia comprar cada música separadamente a custos muito menores por música, lojas de CD's e vinis começam a fechar. Com o avanço do streaming e a popularização de meios como o Spotify, uma outra grande queda na receita, não se paga mais por música ouvida, se paga uma mensalidade que permite escutar uma quantidade praticamente ilimitada de música sem a necessidade de comprar ou possuir a música.

³ Artistas que apenas tem uma música de sucesso.

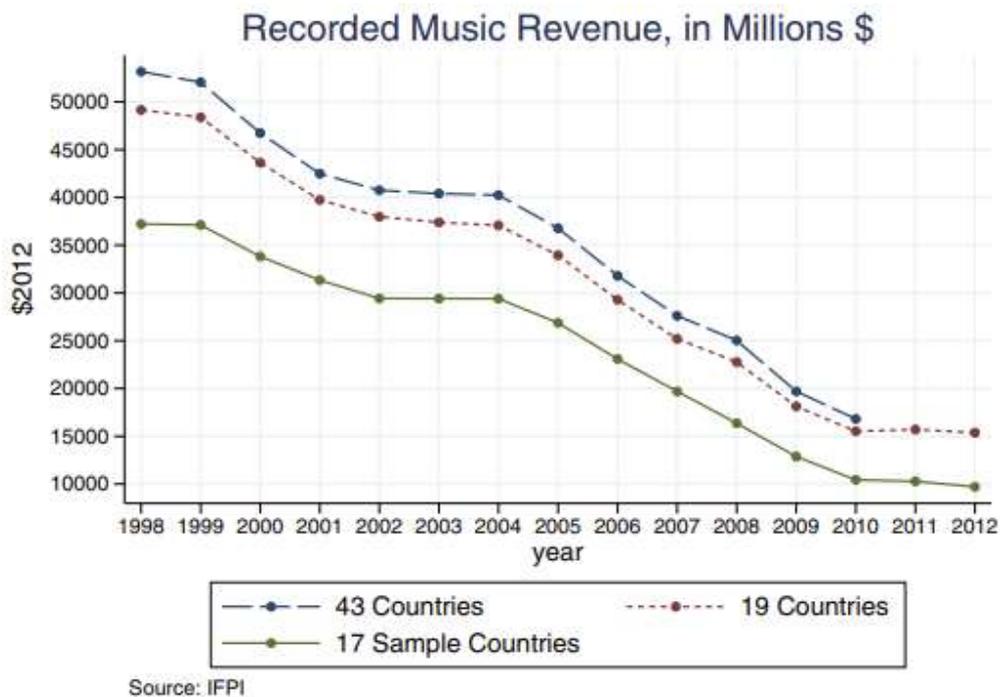


Figura 1

Após um tempo onde já existiam as músicas de forma online e não se fazia mais necessário o consumo da mídia física, a pirataria se tornou um grande problema para o mercado musical, os altos custos envolvidos em ter que comprar cada música eram muitas vezes burlados ilegalmente

C. Information and the Skewness of Music Sales

Um paper escrito por Ken Hendricks e Alan Sorensen foi uma grande motivação ao longo do trabalho. O paper escrito em 2009 precede a moda dos streamings, trabalha com vendas de álbuns físicos ainda. Ao questionar o que faz um artista vender mais, duvidam que apenas a qualidade da música seja um diferencial e através de pesquisas mostram que também depende do conhecimento sobre o artista.

Comparando vendas de álbuns antigos quando artistas lançam álbuns novos, encontram que existe um Spillover. Isso se explica que ao conhecer o artista, consumidores passam a procurar outros álbuns do próprio artista, uma vez que já confiam em seu trabalho. Ao conferir semanas antes do lançamento, quando as músicas ainda não tinham sido lançadas, descobrem que as vendas dos álbuns passados também crescem. Portanto o

trabalho de propaganda do artista é um fator muito importante para o consumidor. Encontram que segundos álbuns vendem 25% a mais do que se não tivessem um álbum lançado antes, o que reafirma a necessidade da indústria de criar contratos longos.

Embora o trabalho deles foque mais em spillovers de novos álbuns em álbuns antigos, o presente trabalho busca um foco maior em spillovers de conhecimento do artista para os trabalhos atuais. Artistas modernos costumam lançar menos álbuns e mais singles. Não há custos em ouvir outras músicas do artista e simplesmente parar de ouvir quando não gostam.

D. Backward Spillover

Um termo comum dentro do cenário econômico tomará um significado muito específico, que a partir de agora usaremos ao longo deste estudo desta forma para evitar maiores complicações.

Definido por Hendricks, K.; Sorensen (2009) “...the effects of new album releases on sales of previous albums by the same artist. The promotional activity and radio airplay associated with a newly released album enhance consumer awareness about the artist and cause some consumers to discover and purchase the artist’s past albums (which are referred to in the industry as “catalog” albums).”

Há evidências do aumento de vendas de álbum, medido desde 13 semanas anteriores e durante anos seguintes, que sugerem que o conhecimento do novo álbum gera um conhecimento do álbum antigo, e do próprio artista.

“During the period covered by our data, radio airplay was the primary channel through which consumers learned about albums. Scarce airtime and the desire of radio stations to get the largest possible audience created an informational bottleneck, with playlists focusing on a small fraction of popular new albums. As a result, sometimes consumers did not buy the right albums or at the right time. The backward spillover reflects consumers correcting initial “mistakes” and buying the right albums at a later time.”

“But in many cases they cannot make these corrections. Artists whose

early albums are mediocre are likely to have their careers truncated. For example, if an artist's first album is only moderately successful, her label may decline to produce any future albums, even though our findings suggest that with full information the artist may have eventually become a success. This of course does not mean that all unsuccessful artists are potential stars. But it does suggest that some potential stars' careers may be truncated because consumers were unaware of their music." (Hendricks, Sorensen 2009)

E. Youtube

Um dos sites mais conhecidos mundialmente começou apenas para vídeos, porém rapidamente se tornou ambiente para ouvintes de música também, hoje conta até com uma plataforma exclusiva para músicas.

F. Tiktok

Uma plataforma que permite usuários gravarem vídeos e colocar músicas de fundo se tornou um dos aplicativos mais populares no mundo. Neste aplicativo, grandes influenciadores costumam gravar vídeos, sejam dançando ou contando alguma história, a isso escolhem uma música para acompanhar. A partir disso a moda mantém a mesma dança e a mesma história atreladas a mesma música, portanto as músicas costumam se repetir junto da moda. Existem várias modas dentro do aplicativo, mas as que mais nos interessam são as que usam apenas a música como áudio dos seus vídeos. Cada moda usa a música de uma maneira diferente, alguns usam para uma dança, outros usam como forma de criar uma história onde o usuário que já conhece a moda, apenas de ouvir o início da música já é capaz de saber quão interessado ele estará no vídeo, uma vez que as músicas costumam, por conta da moda, ficarem atreladas ao mesmo tipo de história já contada previamente por outros usuários.

Tornou-se um grande expoente de música, uma vez que a música dançada por qualquer grande influenciador, muito certamente entrará na moda. Abriu caminhos para novos tipos de marketing e propagandas, antes pagavam para a música tocar no horário nobre da radio, agora pagam para grandes influenciadores dançarem ela.

3. Metodologia e Dados

No passado existiam poucos agentes distribuidores de conhecimento, para se ouvir uma música costumava-se ou ouvir em rádios ou comprar um álbum através de alguma indicação, seja por amigos ou *Charts*⁴. Com o avanço dos meios de comunicação se esperava que os agentes aumentassem e a disponibilidade de todo conhecimento acessível permitisse que mais artistas se tornassem conhecidos. Porém busca-se mostrar que isso não é uma verdade. Embora haja um aumento do número de agentes ainda continua sendo um pequeno nicho de artistas que fazem sucesso. Esses são os artistas que por algum motivo se tornam conhecidos.

Para demonstrarmos isso usaremos dados coletados através da plataforma Chartmetric com uma API desenvolvida especificamente para isto. Os dados sendo principalmente:

- Quantidade de visualizações no Youtube para cada música ao longo do tempo

Com este, usaremos como Proxy para sabermos quanto esta música está sendo ouvida, mesmo que o Youtube seja uma plataforma de vídeos, é muito usada para ouvir músicas e não tem custos atrelados a ele. Portanto, uma música que está sendo ouvida no Youtube é uma música que muito provavelmente está sendo ouvida em outros serviços de streaming.

- Quantidade de vídeo feitos no Tiktok com cada música ao longo do tempo

Usaremos este dado como externalidade informacional para o conhecimento da música. Como o aplicativo do Tiktok não permite ouvir a música toda, não é utilizado para ouvir música, e como as músicas não são escolhidas por serem boas, mas sim por estarem na moda de se dançar ou completar algum desafio enquanto a música toca, sabemos que a qualidade da música pouco importa para os usuários deste aplicativo, podendo assim usarmos como externalidade.

Buscamos encontrar que um aumento de vídeos feitos no Tiktok leve a um aumento de visualizações na plataforma youtube. Ou seja, nossa externalidade que leva a conhecimento aumentando as visualizações dos artistas, em outras palavras, o artista se tornando conhecido por uma externalidade que não é a qualidade da sua música.

⁴ Tabela (paradas) musical criada pela popularidades das músicas

A existência de muitas plataformas de *streaming* e meios de consumos de música cria um espalhamento dos usuários por diversas plataformas. É de se esperar que dependendo da plataforma exista alguma diferença entre os usuários. Por exemplo, Soundcloud, uma plataforma voltada para artistas independentes, onde o usuário faz o próprio upload e não sua distribuidora⁵, terá uma distribuição de ouvintes diferente de outras plataformas mais comuns, como Spotify ou Apple Music. Entre as plataformas mais comuns não existe razão para se esperar ouvintes diferentes entre elas.

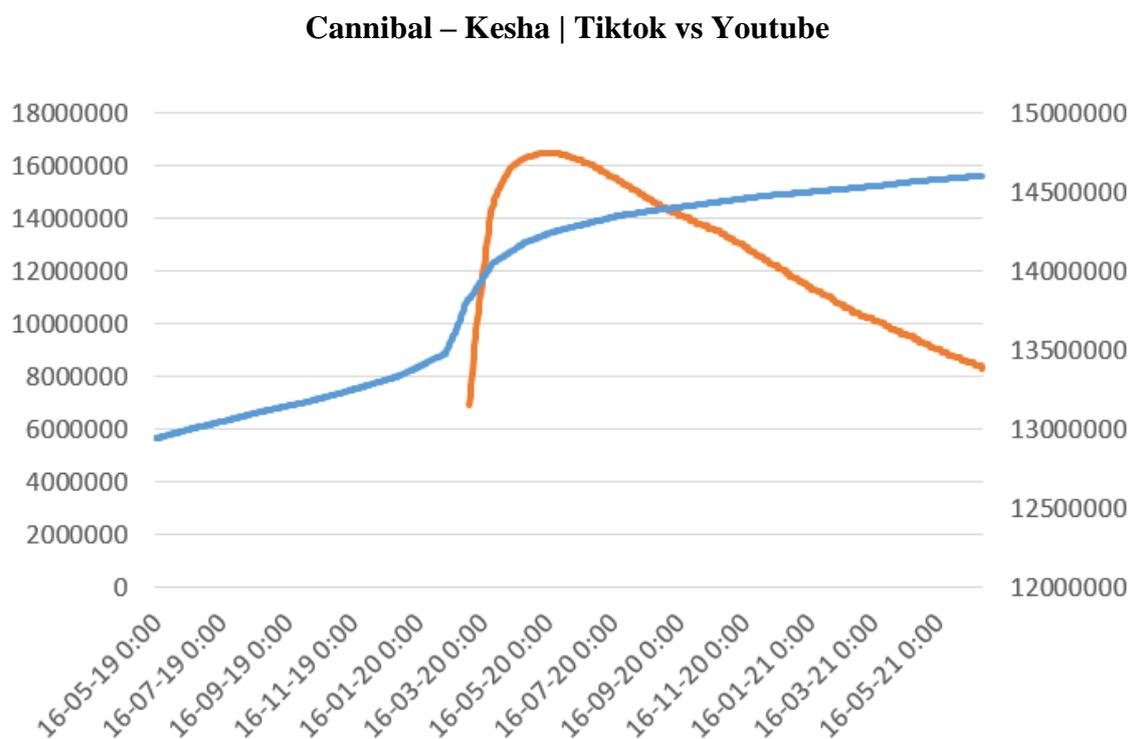


Figura 2

Na figura acima podemos ver como existe uma correlação para as duas curvas. Em laranja vemos a curva de vídeos no Tiktok, em azul o número de visualizações no Youtube. O que pode não ser muito trivial em primeiro momento é enxergar os problemas e consertá-los.

⁵ Agências que fazem o intermédio entre o artista/gravadoras e os meios de venda da música, seja streaming ou vendas físicas. São responsáveis por distribuir a musica as lojas.

A. Seleção de Dados

Selecionar os dados foi uma parte relativamente fácil do trabalho, podíamos trabalhar apenas com dados que viriam a existir, devido a uma dificuldade de coletar dados passados. Portanto não haveria opção de coleta de dados. Outras limitações levaram a escolher três grandes provedores de dados, uma vez que eram os únicos disponíveis: Visualizações no Youtube, vídeos no Tiktok, Shazams ⁶da música. Outra grande limitação é a impossibilidade de adivinhar que músicas se tornariam virais, portanto a coleta de dados só é capaz de começar uma vez que ela já começou a viralizar, veremos este problema e as formas de contorna-lo mais à frente.

Seleciona-se o mês de abril, onde pegam-se as 200 músicas mais populares do Tiktok nos Estados Unidos neste mês. Apenas 123 destas músicas eram utilizáveis, algumas por falta de dados ou por não se encaixarem como uma música propriamente dita, pode apenas ser um discurso ou uma combinação de várias músicas juntas. Este mês é apenas o que definirá que música usaremos, porém, a coleta de dados começou previamente a isto, então teremos dados de acordo com cada música.

B. Problemas e Soluções

O primeiro problema é a curva de número de vídeos do Tiktok diminuindo, embora este ponto de inflexão seja interessante para medirmos quando que a moda de vídeos para de existir e começa a ser interessante deletar os vídeos já postados, pode nos trazer problemas nas contas, já que as visualizações no Youtube apenas são adicionadas, portanto não decrescem. Para consertarmos este problema, todos dados que estiverem decaindo passarão a ser igualados com o último dado positivo.

⁶ Aplicativo para reconhecimento de músicas escutando as músicas.

Cannibal – Kesha | Tiktok vs Youtube [correção I]

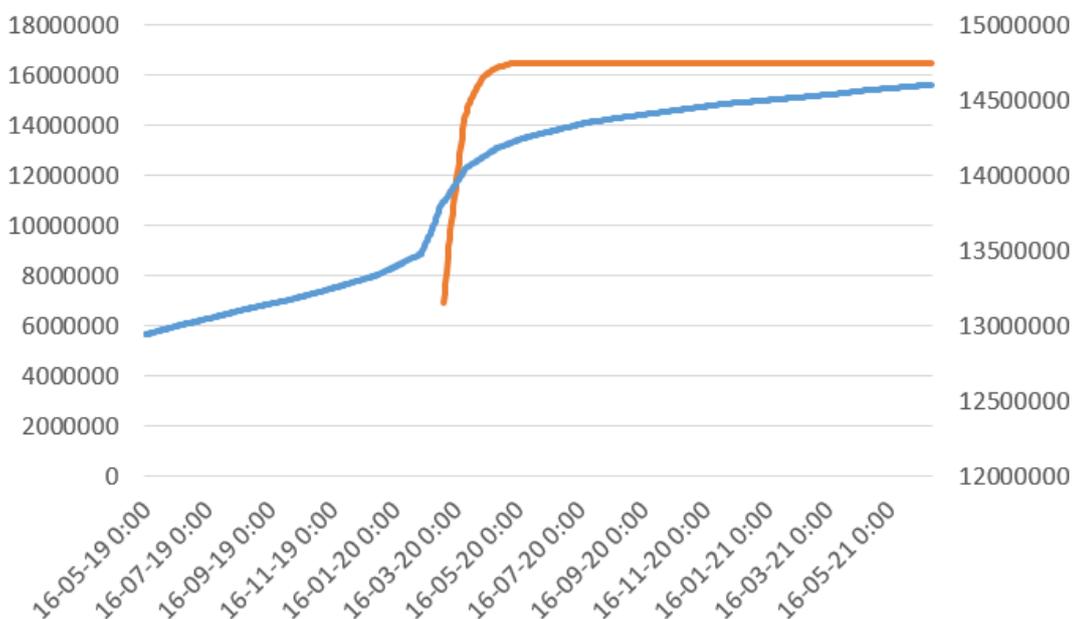


Figura 3

Agora temos um gráfico onde ele foi corrigido para os decréscimos de vídeos do Tiktok. Um novo problema que será muito recorrente também é a falta de dados. Como não temos como prever que músicas serão sucesso no Tiktok e dezenas de milhares de músicas novas são adicionadas, apenas temos acesso a elas depois delas começarem a viralizar, portanto perdemos o começo dos gráficos, não conseguiremos descobrir o tempo de defasagem entre uma música se tornar viral e começar a ser escutada no Youtube, porém podemos imaginar que seja bem curto devido à facilidade moderna.

Ainda se tratando do tratamento dos dados encontramos um problema de ruído, como a coleta deles não foi diária, existem dias com picos muito grandes e dias sem coleta, para suavizar este problema usaremos média móvel de 7 dias, ou seja, médias semanais. Abriremos mão de dados de 6 dias para conseguir isso, porém na grande maioria dos exemplares, estes 6 dias não custam quase nada ao nosso trabalho.

Por último, temos o problema temporal, cada música se tornou viral no seu tempo e data, para isso, trataremos os dados como dia 1 em diante, e não sua data verdadeira. Assim evitaremos um problema de datas aleatórias pra cada artista.

C. Quebras Estruturais

Usaremos um teste de estabilidade econométrica para encontrar quebras estruturais, divide-se em 3 tempos no gráfico:

- Antes da moda no Tiktok;
- Durante a moda no Tiktok, durará desde o início da aceleração até o momento em que os vídeos nos Tiktok se tornem negativos, onde a partir daí será considerado uma defasagem até considerar o fim da moda no Tiktok;
- Depois da moda no Tiktok.

O segundo tempo é muito claro, por análise de gráfico, que se encontrará uma quebra estrutural quando comparado com os outros dois tempos, agora entre o primeiro e o terceiro tempo ele não é tão claramente visto.

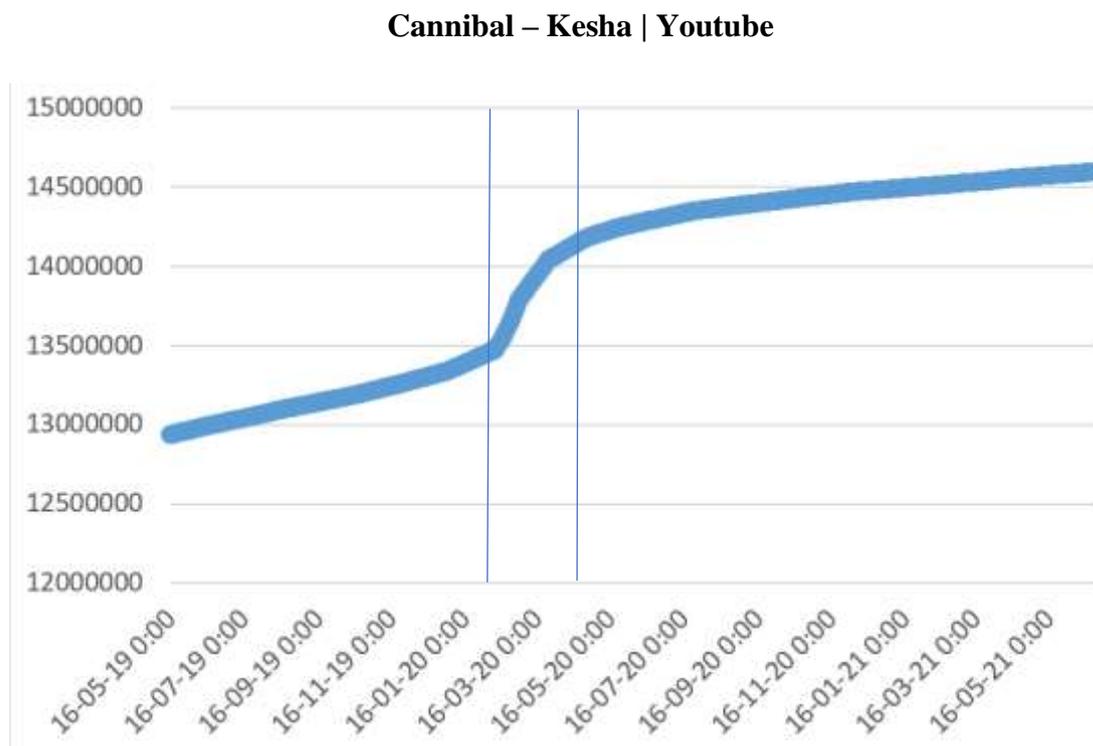


Figura 4

Por exemplo, quando olhado Cannibal da Kesha, uma artista que já era muito conhecida antes do Tiktok, é de se esperar que depois do Tiktok, a fama dela não cresça

significativamente, portanto o tempo 1 e 3 tem variações muito parecidas. Ou seja, a fama adicional que o Tiktok proveu não foi um diferencial para a fama total da artista.

Coño – Puri | Youtube

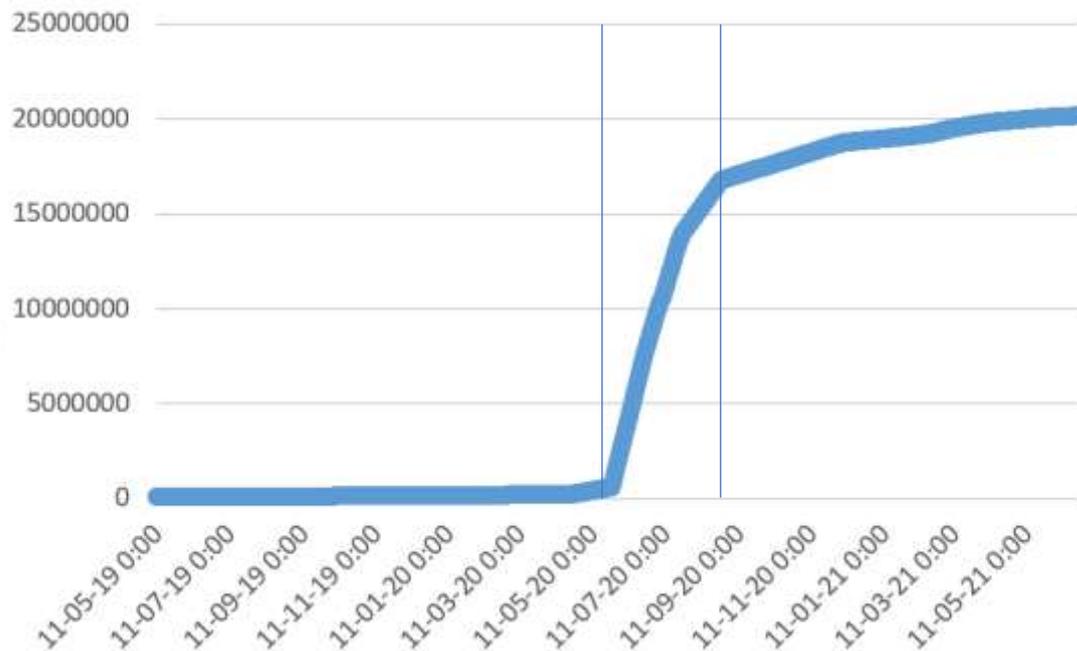


Figura 5

Neste segundo exemplo, quando olhamos Coño, do artista Puri, um artista até então desconhecido, ele goza de uma nova adição de fama provida pelo Tiktok, onde o momento 1 e 3 são diferentes, 1 sendo uma curva praticamente horizontal, onde a falta de fama não permitia acumulo de visualizações e após o fim da moda do Tiktok, que mesmo assim continua acumulando visualizações e assemelhando sua curva com a da artista famosa Kesha, uma vez que agora ele é um artista que sofreu um choque de externalidade para conhecimento.

4. Análise

Nosso Dataframe consiste de 13639 linhas em 27 colunas de datas, reduz-se para 7 colunas, sendo elas: Track_name, Track_duration, Track_artists, Release_date, Track_youtube_views, Tiktok_views (mesmo sendo quantidade de vídeos na plataforma), timestp. Sendo Track_name e Track_artists Character Data e Track_release_date é em formato de data. Agrupa-se os artistas pelas suas músicas, encontrando 82 artistas e músicas utilizáveis.

Embora não seja tão importante, as datas são desde 31 de dezembro de 2020 até 29 de junho de 2021

Análise numérica

```

track_duration_ms track_youtube_views tiktok_view
Min.      : 49084      Min.      :      0      Min.      :      1
1st Qu.  :147350     1st Qu.  :      0      1st Qu.  :  13900
Median   :181500     Median   :      0      Median   : 163600
Mean    :186080     Mean    : 52045172     Mean    : 1905991
3rd Qu.  :220306     3rd Qu.  : 52077751     3rd Qu.  : 1200000
Max.     :504633     Max.     :823668288     Max.     :28600000

```

Tabela 1

Análise numérica II

```

timestp      mean_track_len mean_youtube_views mean_tiktok_views
Min.   :2020-12-31  Min.   :181749  Min.   :43480274  Min.   :1687825
1st Qu.:2021-02-14  1st Qu.:183781  1st Qu.:48749542  1st Qu.:1825154
Median :2021-03-31  Median :186178  Median :52092714  Median :1851818
Mean   :2021-03-31  Mean   :185784  Mean   :52854278  Mean   :1968867
3rd Qu.:2021-05-15  3rd Qu.:187710  3rd Qu.:53732766  3rd Qu.:1985972
Max.   :2021-06-29  Max.   :187935  Max.   :69586338  Max.   :2777859

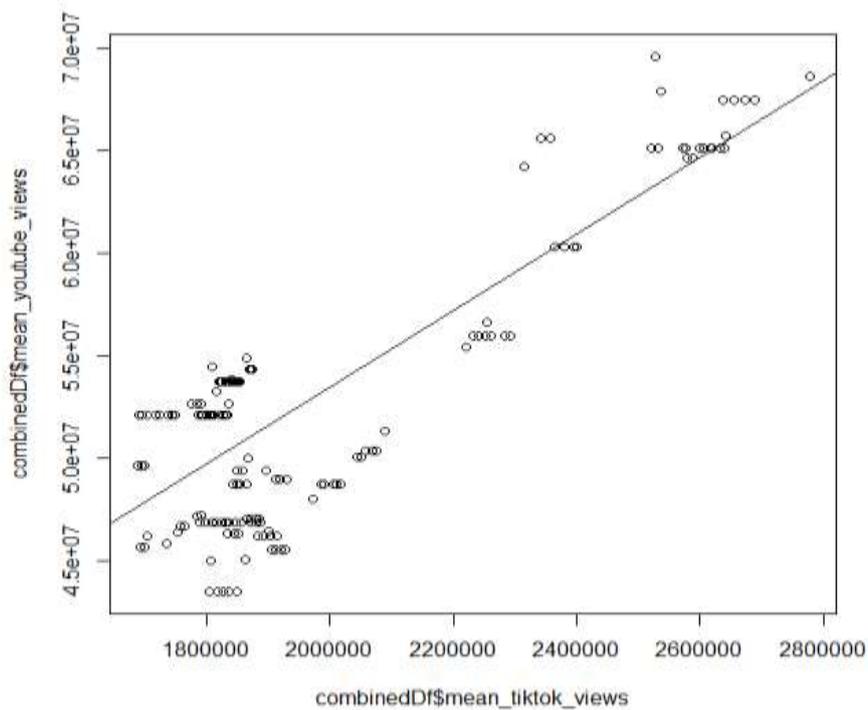
```

Tabela 2

A. Análise de Correlação

Investigaremos se existe alguma correlação entre duas variáveis independentes, em outras palavras: como a variação em uma das variáveis faz a segunda variável se comportar.

Linha ab entre Numero de visualizações no Youtube e quantidade de vídeos no Tiktok



Da figura acima consegue-se enxergar uma correlação positiva, a quantidade de vídeos no Tiktok é positivamente relacionada com a quantidade de visualizações no Youtube. Isso já nos dá uma primeira dica sobre a relação que procuramos, porém precisamos ir mais fundo no assunto, uma vez que queremos mostrar que a quantidade de vídeos que influencia a quantidade de visualizações e não o caminho contrário.

Coeficiente de correlação de postos de Spearman nos permitirá testar ainda mais profundamente a correlação entre essas duas variáveis, os valores para o teste podem variar entre 1 e -1 para o coeficiente de correlação de Pearson (ρ_0) e para ele faremos um teste de hipótese:

$H_0: p_s = 0$: Youtube e Tiktok não são relacionados

$H_A: p_s \neq 0$: Youtube e Tiktok são relacionados

Para obter o coeficiente procurado, r_s é calculado para o teste. Primeiro é necessário colocar as visualizações em ordem crescente e suas respectivas observações de números de vídeos no Tiktok. A diferença entre os rankings binários de observação é calculada por:

$$\sum d_i = 0$$

E a diferença do quadrado é: $\sum d_i^2 = 517220$

O coeficiente de correlação de postos de Spearman é calculado por:

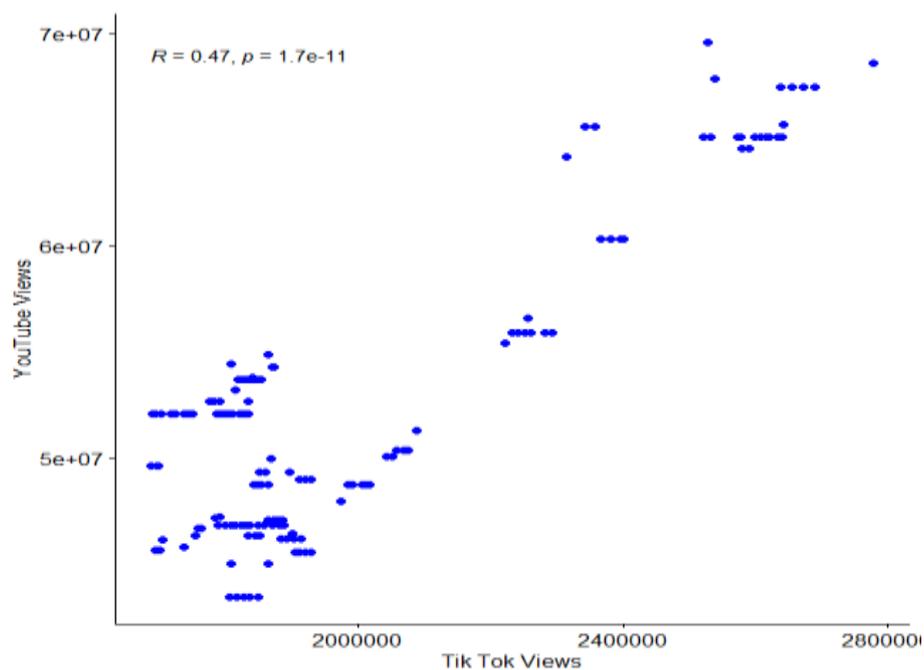
$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

Onde encontramos $r_s = 0.48$.

O coeficiente diferente de zero e positivo afirma a hipótese alternativa, não é capaz de afirmar a hipótese nula

Observando o nível de significância de 0.05 para conferir H_A podemos fazer um gráfico de dispersão

Gráfico de dispersão



O coeficiente de correlação para postos de Spearman ($R=0.47$) é o mesmo que o de cima. Seu P-valor é menor que 0.05 para a significância estatística, o que significa que a hipótese alternativa é verdadeira ($H_A: \rho_s \neq 0$).

O ponto interessante que se traz à tona é o que o Tiktok diferencia na vida de um artista, e por Tiktok queremos dizer qualquer tipo de externalidade que faça o artista mais famoso, mais conhecido. Enquanto os números de Tiktok aumentam, suas visualizações aumentam, levando o artista a ter fama, porém uma vez que a moda do Tiktok passa e para-se de se usar a música no aplicativo, para artistas que já tinham a música conhecida, voltam ao normal, porém para artistas que sua música era inédita ao público, a volta pós a moda tem uma inclinação maior que antes, tendo assim, feito uma diferença no conhecimento do artista.

5. Conclusão

Mostrou-se que uma música se tornar viral no Tiktok gera uma grande influência na performance da música em outra plataforma. Com isso, um bom caminho para provar que o conhecimento do artista leva a uma melhor performance da sua música e como provado por Hendricks e Sorensen (2009) o conhecimento sobre o artista é uma, se não a principal forma de promover os futuros trabalhos dos artistas.

Contas adicionais e mais dados poderiam permitir cálculos mais profundos de como a carreira geral do artista poderia ser beneficiada pelo Tiktok ou outras externalidades, como por exemplo, quantas vezes é útil a um artista tentar viralizar no Tiktok? E ao longo do tempo poderia se perguntar o prazo de memória para artistas.

Neste caso pergunta-se: viralizar algumas vezes em pouco espaço de tempo é tão, menos ou mais eficiente que viralizar algumas vezes em um espaço de tempo maior? Porém o aplicativo ainda é recente demais para podermos ter essa coleta de dados tão extensa.

No entanto, demos um primeiro passo a mostrar que a externalidade de um aplicativo leva a outros meios de streaming seguirem sua externalidade, independente da qualidade da música, já que sabemos que no Tiktok existem muitas músicas que viralizam que poderiam ser consideradas péssimas, caso isso fosse um estudo sobre qualidade.

6. Referências Bibliográficas

Hendricks, K.; Sorensen, Information and the Skewness of Music Sales Journal of Political Economy. Journal of Political Economy, 2009.

Goeree, M. S., Advertising in the U.S. Personal Computer Industry, Working paper, Claremont McKenna Coll, 2005

Dolan, P., & Galizzi, M. M., *Because I'm Worth It: A Lab-Field Experiment on the Spillover Effects of Incentives in Health*, CEP Discussion Papers, 2014

Dolan, P., & Galizzi, M. M., Like ripples on a pond: Behavioral spillovers and their implications for research and policy, *Journal of Economic Psychology*, 47, 1–16, 2015

Ward, Y.; Graves, A., A New Cost Management & Accounting Approach For Lean Enterprises, Working Paper Series, School of Management, 2004

Salganik, M; Sheridan D, P; Watts, D., Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market, *Science* vol. 311 Issue 5762, 2006

Liebowitz, J., measuring the value of online communities, leading to innovation and learning, *International Journal of Innovation and Learning* January, 2003

Shapiro, C.; Varian, H. R., The Art of Standards Wars, *California Management Review*, 1999;41(2):8-32.

Aguiar, L.; Duch-Brown, N.; Waldfogel, J., Revenue, New Products, and the Evolution of Music Quality Since Napster. IPTS Working Paper. IPTS.L. 2014.

Anderson, C., *The Long Tail: Why the Future of Business is Selling Less of More.* Hyperion, 2006

M., Krueger, A.B., Rockonomics: the economics of popular music. In: *Handbook of the Economics of Art and Culture*, vol. 1. Elsevier, 2006, pp. 667–719. (Chapter 20)

Knopper, S., *Appetite for Self-Destruction: The Spectacular Crash of the Record Industry in the Digital Age.* Free Press, 2009.

Liebowitz, S.J., How much of the decline in sound recording sales is due to file-sharing? *J. Cult. Econ.* 2014, 40 (1), 13–28.

Terviö, M., Superstars and mediocrities: market failure in the discovery of talent. *Rev. Econ. Stud.* 76 (2), 2009, 829–850.

Waldfoegel, J., Copyright protection, technological change, and the quality of new products: evidence from recorded music since Napster, 2012, *J. Law Econ.* 55 (4), 715–740.

Waldfoegel, J., Digitization and the quality of new media products: the case of music. In: *Economics of Digitization*. University of Chicago Press, 2013.

Wooldridge, J. M.; *Econometric Analysis of Cross Section and Panel Data*, Massachusetts Institute of Technology, 2010.

Chow, G. C.; Tests of Equality Between Sets of Coefficients in Two Linear Regressions . *Econometrica.* 28 (3): 591–605. doi:10.2307/1910133. JSTOR 1910133. Archived from the original, 2019.

Doran, H. E.; *Applied Regression Analysis in Econometrics*, 1989.

Wooldridge, J. M.; *Introduction to Econometrics: A Modern Approach (Fourth ed.)*. Mason: South-Western. pp. 243–246. ISBN 978-0-324-66054-8., 2009

Hsiao, C.; *Analysis of Panel Data (Second ed.)*. New York: Cambridge University Press. ISBN 0-521-52271-4, 2003.

Spearman, C.; The Proof and Measurement of Association between Two Things. *The American Journal of Psychology.* 15 (1): 72–101. doi:10.2307/1412159, 1904.

Yule, G. U., Kendall, M.; *An Introduction to the Theory of Statistics*. G. Udny Yule, ... and M.G. Kendall, ... 14th Edition Revised and Enlarged (em inglês). [S.l.]: C. Griffin, 1950.

Caruso, J. C.; Cliff, N.; Empirical Size, Coverage, and Power of Confidence Intervals for Spearman's Rho. *Educational and Psychological Measurement (em inglês).* 57 (4): 637–654. doi:10.1177/0013164497057004009, 2016.

7. Apêndice

Formato da análise

```
[1] "2021-06-20"
> dfMusic <- MusicDF %>% select(track_name, track_duration_ms, track_artists, track_release_date, track_youtube_views, tiktok_view, timestp)
> head(dfMusic)
  track_name track_duration_ms track_artists track_release_date track_youtube_views tiktok_view timestp
1 Heartbreak Anniversary      198370      Giveon      2020-03-27             NA      345100 16/02/2021
2 Heartbreak Anniversary      198370      Giveon      2020-03-27      446600 17/02/2021
3 Heartbreak Anniversary      198370      Giveon      2020-03-27             NA      521900 18/02/2021
4 Heartbreak Anniversary      198370      Giveon      2020-03-27             NA      592700 19/02/2021
5 Heartbreak Anniversary      198370      Giveon      2020-03-27             NA      658500 20/02/2021
6 Heartbreak Anniversary      198370      Giveon      2020-03-27             NA      721000 21/02/2021
```

Código R

```
library(readr)
library(tidyverse)
library(dplyr)
library(lubridate)
```

```
music_stats <- read_csv("final_merged_data_with_stats.csv")
#make copy of dataframe
MusicDF<-data.frame(music_stats)
```

```
str(MusicDF)
```

```
head(MusicDF$track_release_date)
head(MusicDF$timestp)
#Convert "track_release_date" and "timestp" to date time object
class(MusicDF$timestp) #Check Class of "timestp"
class(MusicDF$track_release_date) #Check Class of "track_release_date"
```

```
MusicDF$timestp <- as.Date(MusicDF$timestp, format = '%d/%m/%Y')
#####Data Exploration and Cleaning
dim(MusicDF) #Check Size Dataframe
```

```
dfMusic <- MusicDF %>% select(track_name, track_duration_ms, track_artists,
track_release_date, track_youtube_views, tiktok_view, timestp)
```

```
head(dfMusic)
```

```
dim(dfMusic)
```

```
sapply(dfMusic, function(x) sum(is.na(x))) #Count Missing Values
```

```
#dfMusic[is.na(dfMusic)] <- 0 #Impute Missing Values with Zero
```

```
sapply(dfMusic, function(x) sum(is.na(x))) #Count Missing Values
```

```
#Convert "timestp" col to datetime obj
```

```
dfMusic$timestp <- as.Date(dfMusic$timestp, format = '%d/%m/%Y')
```

```
str(dfMusic) #Check datatypes
```

```
#Group Data by Artist
```

```
dfMusic %>%
```

```
  group_by(track_artists) %>%
```

```
  summarise(n_distinct(track_name))
```

```
min(dfMusic$timestp)
```

```
max(dfMusic$timestp)
```

```
#Numeric Data
```

```
numericDf <- dfMusic %>% dplyr::select(where(is.numeric))
```

```
summary(numericDf)
```

```
#Convert Character columns to factor
```

```
#dfMusic$track_name <- as.factor(dfMusic$track_name)
```

```
#dfMusic$track_artists <- as.factor(dfMusic$track_artists)|
```

```
#Categorical variables
```

```
summary(dfMusic$track_artists)
```

```
##Count Unique
```

```
##Create dataframe for containing mean Views for Youtube and Tik Tok per day
```

```
dfMusic_yT <- dfMusic %>%
```

```
  mutate(timestp = floor_date(timestp)) %>%
```

```
  group_by(timestp ) %>%
```

```
  summarize(mean_youtube_views = mean(track_youtube_views))
```

```

dfMusic_TT <- dfMusic %>%
  mutate(timestp = floor_date(timestp)) %>%
  group_by(timestp ) %>%
  summarize(mean_tiktok_views = mean(tiktok_view))

dfMusic_Len <- dfMusic %>%
  mutate(timestp = floor_date(timestp)) %>%
  group_by(timestp ) %>%
  summarize(mean_track_len = mean(track_duration_ms))

#Combine the datasets

combinedDf <- merge(dfMusic_yT, dfMusic_TT, by='timestp')
combinedDf <- merge(dfMusic_Len, combinedDf, by='timestp')

#Analysis of Dataset
#Convert dataframe to time series data
head(combinedDf)
#Summary for combinedDf data
summary(combinedDf)

#Plot Youtube Views across the period
ggplot(combinedDf, aes(x=timestp, y=mean_youtube_views)) +
  geom_line(color="blue")+
  ggtitle("YouTube Views from Dec 2020 to June 2021") +
  xlab("Date") + ylab("YouTube Views")

#Plot Tik Tok Views across the period
ggplot(combinedDf, aes(x=timestp, y=mean_tiktok_views)) +
  geom_line(color="green")+
  ggtitle("Tic Tok Views from Dec 2020 to June 2021") +
  xlab("Date") + ylab("Tik Tok Views")

# Create the scatter plot with `Tick Tok` on the x-axis and `Youtube Views` on the y-
axis
plot(combinedDf$mean_tiktok_views, combinedDf$mean_youtube_views)
# Add a regression line with the form `abline(lm(y ~ x))`
abline(lm(combinedDf$mean_youtube_views ~ combinedDf$mean_tiktok_views))

```

```
df_spearman<- combinedDf %>% mutate(
  rank_TY_view=rank(mean_youtube_views),
  rank_TT_view=rank(mean_tiktok_views),
  d=rank_TY_view-rank_TT_view,
  d_square=d^2) %>%
  select(-mean_track_len)
head(df_spearman)

sum(df_spearman$d)

d_square_sum <- sum(df_spearman$d_square)
d_square_sum

n <- nrow(combinedDf)
rho_s <- (1-(6*(sum(d_square_sum)))/(n*(n^2-1))) %>% round(2)

rho_s

library("ggpubr")

ggscatter(combinedDf, x = "mean_tiktok_views", y = "mean_youtube_views",
  color = "blue", cor.coef = TRUE,
  cor.method = "spearman",
  xlab = "Tik Tok Views", ylab = "YouTube Views")
```