

情報可視化ソフトウェア Hidden によるデータ分析

YouTuberのチャンネル登録者数・総視聴回数
YouTubeにアップされている動画の視聴回数

はじめに

わたしはYouTube依存症と言っても過言ではないくらい、YouTubeを視聴することに人より非常に多くの時間を割いている。YouTubeには膨大な数の動画がアップされており、飽きてしまうことはない。この課題のテーマを考えながら、ふとYouTubeを見ようと携帯に手を伸ばした時、こんなにも好きなYouTubeを徹底的に分析しようと思いついた。YouTuberにとって、チャンネル登録者数、及び総視聴回数は広告収入に影響するため大事な数字であり、視聴者側からすると人気を知ることができるものとなっている。これらは何に影響を受けているのだろうか。また、同じYouTuberがアップした動画でも再生回数にかなりばらつきが見られる場合がある。内容の違いで動画の伸び率が変わることは間違いないのだが、それ以外に影響を及ぼす要素はないのだろうか。

YouTubeの視聴回数とYouTuberのチャンネル登録者数について可視化していこうと思う。

YouTube

- Googleが提供する **世界最大の動画共有サービス**
- サービスは基本**無料**で利用でき、視聴は登録をしていなくても可能である
- 評価をしたりコメントを残したりするには、**無料の会員登録**が必要である
- チャンネルを開設すれば、動画を**自由に投稿**できる

(参考 : <https://ja.wikipedia.org/wiki/YouTube>)

YouTuber

- YouTuberとは自分で作った動画をYouTubeに投稿すること
- チャンネルを開設さえすれば、誰でもなることができる
- チャンネル登録者数や総再生時間などの一定の基準をクリアすると、広告収入を得ることができる
- 広告収入だけで生活しているYouTuberもいる
- 正式な「職業」として認知されるようになった

(参考：<https://ja.wikipedia.org/wiki/YouTuber>)

YouTubeのユーザー層



- どの世代も利用率は高い
- 利用率が特に高いのは10代である
- どの世代も、男女のユーザー割合に大きな差はない
- 最も多いのは40代、次に多いのが50代、その次が30代である

(出典： <https://note.natee.jp/n/n2e18746fe99f>)

各SNSの特徴

• Instagram

写真共有SNS。多くのフォロワーを持ち強い影響力を持つインフルエンサーは、**インスタグラマー**と呼ばれ、大きな広告効果を持つ。24時間だけ短い動画や画像が投稿される**ストーリー機能**がある。

• Twitter

テキスト、動画、画像、URLを投稿できるSNS。他の人の投稿を再投稿する**リツイート機能**や、多く投稿されている単語がわかる**トレンド機能**がある。

• TikTok

ショートビデオ共有SNS。BGMをつけたり、動画を簡単に編集したりすることができる。アップした動画は保存することもできる。

(参考: <https://ja.wikipedia.org/wiki/Instagram>、 <https://ja.wikipedia.org/wiki/Twitter>、 <https://ja.wikipedia.org/wiki/TikTok>)

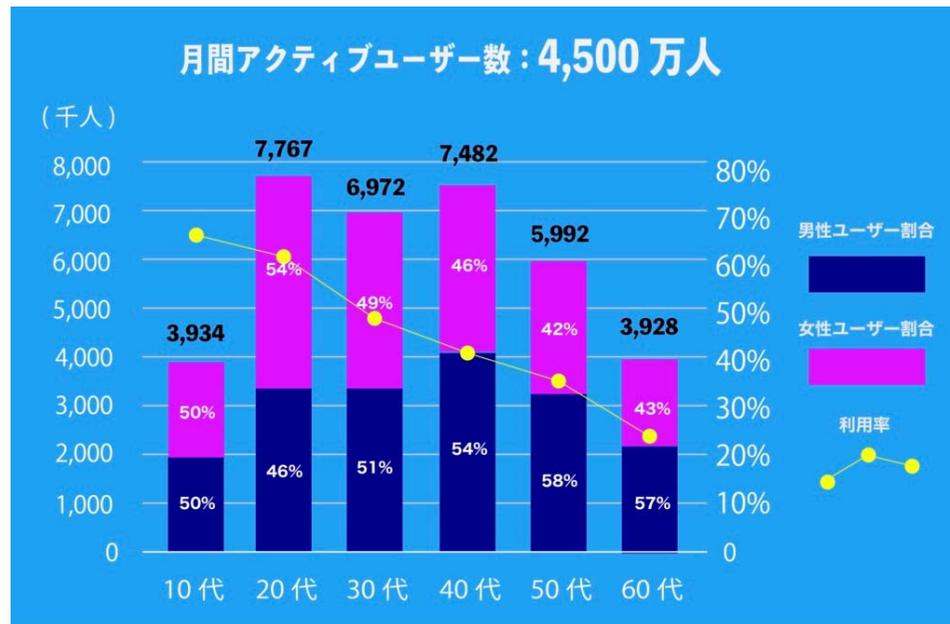
Instagramのユーザー層



- 利用率は10代が最も高く、20代、30代…とどんどん減っている
- 20代から40代の利用者が多い
- 利用者が最も多いのは20代である
- 20代から40代は女性ユーザー割合が高く、50代から60代は男性ユーザー割合が高い

(出典： <https://note.natee.jp/n/n2e18746fe99f0>)

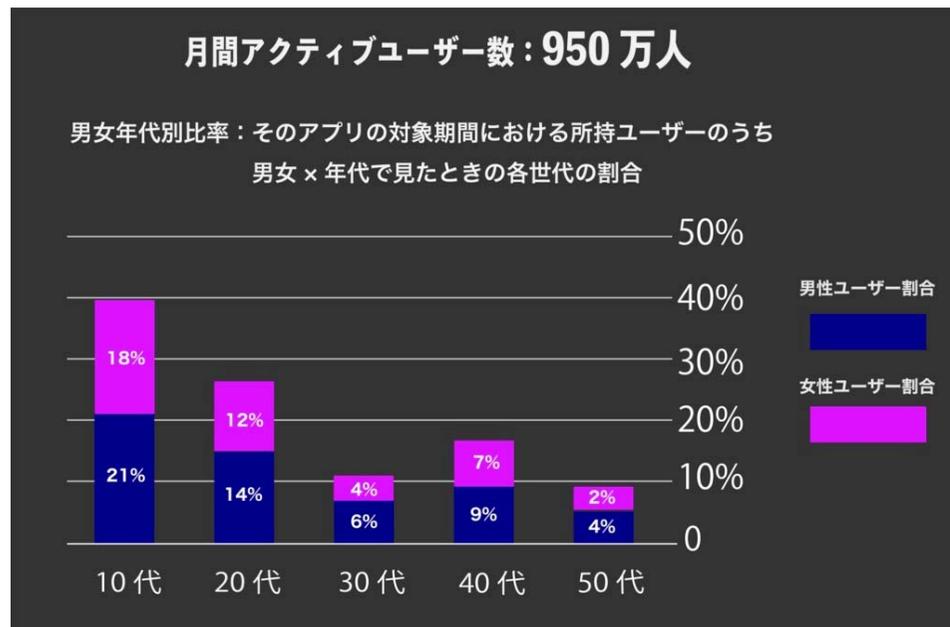
Twitterのユーザー層



- 利用率は10代が最も高く、20代、30代…とどんどん減っている
- 20代と40代の利用者が多い
- 10代と60代の利用者は少ない
- 男性ユーザー割合の方が高い世代が多い
- 全体的にInstagramより利用者が多い

(出典： <https://note.natee.jp/n/n2e18746fe99f>)

TikTokのユーザー層



- 10代、20代の利用者が全体の約7割をしめている
- 30代から50代の利用者は10代、20代の利用者に比べるとかなり少ない
- 男性ユーザー割合と女性ユーザー割合に大きな差がある世代はない

(出典： <https://note.natee.jp/n/n2e18746fe99f>)

使用するデータ①

○YouTuberに関する可視化

- ーチャンネル登録者数
- ーチャンネル開設日
- ー動画の本数
- ー投稿頻度(本/年)
- ー最高再生回数
- ー総視聴回数
- ーYouTuberの人数

ーサブチャンの有無

ーメンバーシップの有無

ー事務所の有無

0 : 無し 1 : 有り

ー性別

0 : 男性 1 : 女性 2 : 男女

ー芸能人かどうか

0 : 芸能人以外

1 : 芸能人

(これらのデータはYouTubeを参照した)

使用するデータ①

- ーInstagramのフォロワー数
1000：アカウント無し
- ーTwitterのフォロワー数
5000：アカウント無し
- ーTikTokのフォロワー数
1000：アカウント無し

(これらのデータはInstagram、Twitter、TikTokを参照した)

項目選定の意図

★チャンネル登録者数

チャンネル登録をしている視聴者の数。YouTubeは視聴者がチャンネル登録をすると、そのチャンネルが新しい動画をアップするたびに通知が来る機能があり、さらに、視聴者の画面上には登録チャンネルのみのタイムラインの欄がある。チャンネル登録者数が多ければ多いほど動画をほぼコンスタントに見る視聴者が多いと言える。つまり、チャンネル登録者数はYouTuberの**知名度**や**人気度**を図る指標でもあると言える。

★チャンネル開設日

チャンネルを開設した日から現在までの期間の長さはファンを獲得することができた期間の長さと言える。YouTuberに対して「推し」という概念を持つ人も多く、長い期間YouTuberをやっている人は新しいファン（新規）獲得に加えて古くからのファン（古参）も多いと思われる。ファン側としては応援している期間が長ければ長いほど応援しようとする気持ちが強く、離れにくいファン（固定ファン）が多くいるのではないかと思う。

- ー推し：イチ押しのメンバーを意味する略語“推しメン”をさらに略したもの
- ー新規：もともとファンではなかったが、新たにファンになった人
- ー古参：昔から応援し続けているファン

(参考：<https://numan.tokyo/words>、

項目選定の意図

★動画の本数

アップしている動画の本数。動画の本数が多いほど、YouTuberのコンテンツが増えることになり、視聴者の検索に引っかかる可能性が高くなる。また、視聴者としては動画の本数が多い方が飽きずにそのYouTuberの動画を見ることができると思う。

★投稿頻度(本/年)

一年あたりにアップしている動画の本数。投稿頻度が高いと、動画の本数が多くなるためファンは飽きずらくなり更新を楽しみにするが、内容が薄くなったり、編集に時間がかけられなくなったりする。一方で、投稿頻度が低いと、内容は濃くなるため同じような動画が複数あることを防げたり、編集に時間をかけることができるが、動画がたくさん上がる方がファンにとっては嬉しいため、ファンは物足りなさを感じる。

★最高再生回数

YouTuberがアップした動画の中で最も再生された動画の回数。動画の再生回数が爆発的に伸びることを、「バズる」という。一度バズった動画をきっかけにチャンネル登録者数が格段に増える場合がある。また、総視聴回数に大きく貢献する。

項目選定の意図

★総視聴回数

総視聴回数とは、全動画の再生回数の合計回数のことである。総視聴回数が多いということは、アップしている動画のほぼ全部がある程度の再生回数を保持している、またはバズった動画があるということである。

★YouTuberの人数

グループで活動しているYouTuberはグループ全体のファンもいるが、グループ内の特定のメンバーのファンもいると思う。グループの人数が多ければ多いほど獲得できるファン層が広がるのではないかと思う。しかし、YouTuberの人数が多いと、全ての動画に出演するわけではなく出る動画を分担しているYouTuberも多くいるので、特定のメンバーに対するファンはそのメンバーが出ていない動画は見ない可能性もある。したがって、メンバーが多ければ多いほどチャンネル登録者数が増える可能性はあっても、総視聴回数が伸びるとは一概にも言えないと思う。

★サブチャンの有無

メインのチャンネルに加えて他に持っているチャンネル。メインチャンネルとは違う雰囲気動画を出したり、メインチャンネルに比べて緩い動画を出したりするYouTuberが多い。先にサブチャンでYouTuberを知り、メインチャンネルのチャンネル登録に繋がる可能性がある。

項目選定の意図

★メンバーシップの有無

YouTube上での、YouTuberのファンクラブのようなもの。月額制で、メンバーシップ限定の動画がアップされたり、生配信が行われる。チャンネル登録者数が多いYouTuberの方がメンバーシップを設けているイメージがある。

★事務所の有無

事務所に入っているかどうかである。無所属で活動するYouTuberも多い一方、クリエイターやインフルエンサーとして事務所に所属するYouTuberもいる。事務所に所属することで、案件を取りやすくなったり、YouTuberとして活動する上でのサポートを受けることができたりする。しかし、いろいろな制約が加わったり、許可が降りないことがあったりするので、自由度は減る。

★性別

YouTubeは男女のユーザー割合に大きな差がないため、チャンネル登録者数に大きな影響は及ぼさないのではないかと思う。男女で動画のリピート率に差があれば、総視聴回数に差が出る可能性がある。

★芸能人かどうか

芸能人は元々の知名度があるため、チャンネル登録者数が伸びやすい。しかしテレビ受けするコンテンツと、YouTube受けするコンテンツは違う可能性がある。

項目選定の意図

★Instagramのフォロワー数

Instagramは20代から30代の利用者が多いのに対して、YouTubeの利用者は30代から50代が中心であり、少し年齢層はずれている。しかし、フォロワーの多いインフルエンサーはたとえ一般人でも大きな発信力を持つため、Instagramのフォロワー数とチャンネル登録者数は関係があると思われる。

★Twitterのフォロワー数

TwitterはInstagramやTikTokに比べて全体の利用者が多い。したがってTwitterはInstagramやTikTokよりも大きく影響を及ぼすと思う。20代から30代に利用者がおおく、YouTubeの利用者層とほぼかぶっている。

★TikTokのフォロワー数

TikTokは10代や20代の若者が中心に利用しており、YouTubeに多い利用者層とは違う。しかし、動画をアップするという点がYouTubeと共通している。

予想(YouTuberに関する可視化)

- 動画の本数が多い方がYouTuberのコンテンツが増えるため、チャンネル登録者は多い
- 投稿頻度が高い方が固定ファンを作りやすいため、チャンネル登録者数は多い
- 各SNSのフォロワー数が多い方がチャンネル登録者数は多い
- 総視聴回数とチャンネル登録者数は比例する
- サブチャンがあるYouTuberはメインチャンネルで視聴回数を保っているということなので、総視聴回数が多い

使用するデータ②

○動画に関する可視化

- 動画の再生回数(number)
- 動画の長さ(length)
- 動画の投稿日(days)
- 高評価数(good)
- 低評価数(bad)
- コメント数(comment)
- ハッシュタグ数(#)

(これらのデータはYouTubeから参照した)

- YouTuberの人数(member)
- コラボ動画かどうか(collaboration)
 - 0 : コラボしていない
 - 1 : コラボしている
- 芸能人かどうか(talent)
 - 0 : 芸能人以外
 - 1 : 芸能人
- 性別(gender)
 - 0 : 男性 1 : 女性 2 : 男女

項目選定の意図

★動画の長さ

YouTubeはテレビと違って早送りをしたり、観たい場所から観たり、違う動画にすぐ飛んだりすることが可能である。したがって、YouTubeに適した動画の長さがあると思う。ただ、動画が10分以上になると広告収入が増えるので、10分を超える動画が多いと思う。

★動画の投稿日

視聴者が動画をリポートすれば、動画が投稿されてから時間が経てば経つほど、動画の再生回数は多くなる。

★高評価数

YouTubeは視聴者が動画に高評価・低評価をつけることができる。その数を視聴者側は見ることができ、高評価が多い動画の方が見られやすい傾向にあると思う。

★低評価数

低評価が多い動画は観るのを敬遠されると思うが、炎上している動画は低評価が多いので、逆に再生回数が多い可能性がある。

★コメント数

視聴者は動画に対して感想を自由に投稿することができる。

項目選定の意図

★ハッシュタグ数

タグ付けすることで検索に引っかかりやすくなり、動画を見つけてもらう可能性が高くなる。

★YouTuberの人数

グループで活動しているYouTuberはグループ全体のファンもいるが、グループ内の特定のメンバーのファンもいると思う。グループの人数が多ければ多いほど獲得できるファン層が広がり、再生数に繋がると思う。

★コラボ動画かどうか

他のYouTuberとコラボする動画がある。チャンネル登録していない視聴者も見ることがあるので、再生数が伸びることに繋がる。

★芸能人かどうか

元々の知名度があるため、チャンネル登録をしていなくてもおすすめに表示された時に再生する可能性がある。

★性別

男女で動画のコンテンツが違うため、動画のリピート率が変わると思う。例えばメイク動画は、視聴者が参考にメイクをする時に何度も再生されるため、再生回数が伸びる。

予想(YouTubeに関する可視化)

- 芸能人は元々の知名度があるので、再生回数が伸びる
- 高評価数が多い動画は再生回数が伸びる
- コラボ動画は再生回数が伸びる
- ハッシュタグ数が多い方が、検索に引っかかりやすくなるため、再生回数が伸びる
- 動画の投稿日が早ければ早いほど、視聴者の再生する期間が長くなり、さらにリピート再生する視聴者がいるため、再生回数は伸びる

散布図

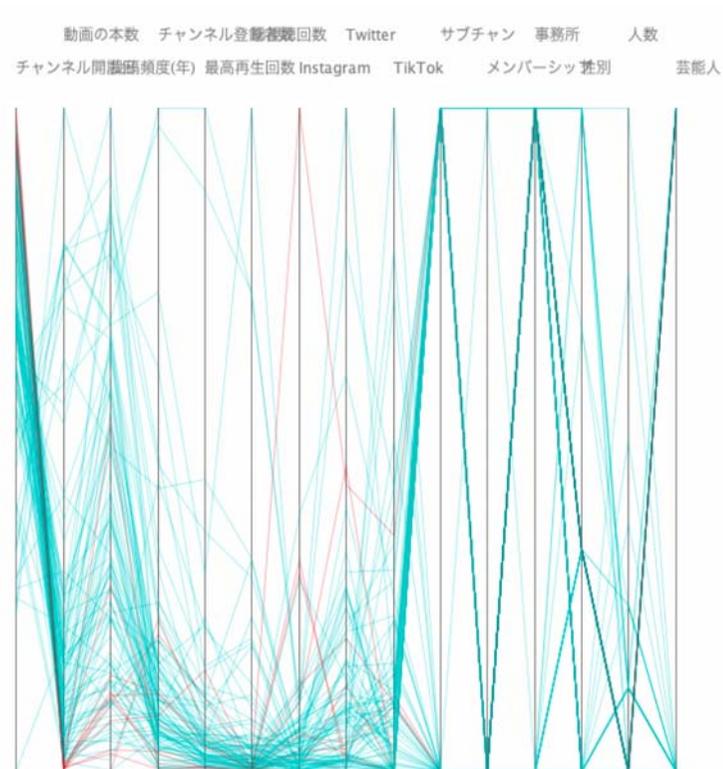


Instagram、Twitter、TikTokはチャンネル登録者数と相関があると予想していた。点が離れているのは意外であった。

チャンネル登録者数と総視聴回数は相関が強いことが分かる。

また、総視聴回数と動画の本数も相関が強いことが分かる。

クラスタリング



○Clustering=2

ほとんどの線が水色である。

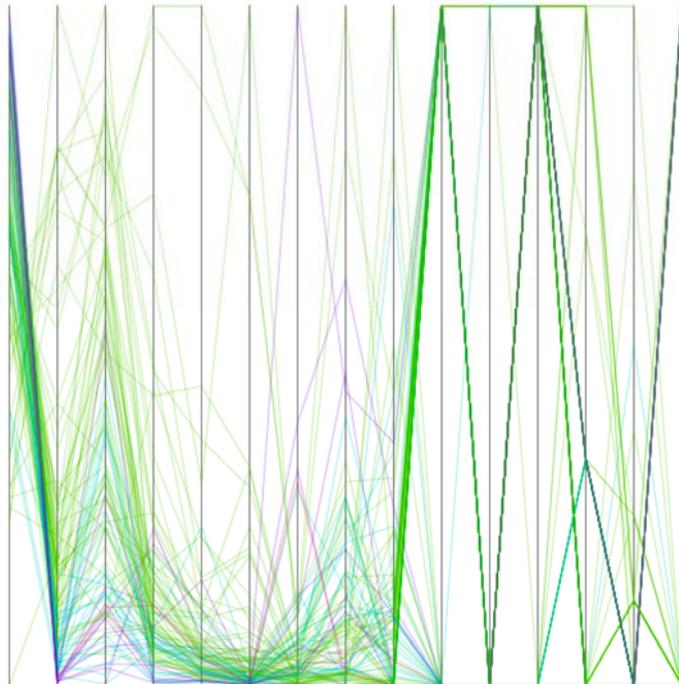
赤の線は数本しかない。

「動画の本数」の軸と「総視聴回数」に注目する。動画の本数が少なく、かつ、総視聴回数の少ない線が赤であることがわかる。

水色の線は、単にそれ以外の動画のことを表している。

クラスタリング

動画の本数 チャンネル登録視聴回数 Twitter サブチャン 事務所 人数
チャンネル開設頻度(年) 最高再生回数 Instagram TikTok メンバーシップ 芸能人



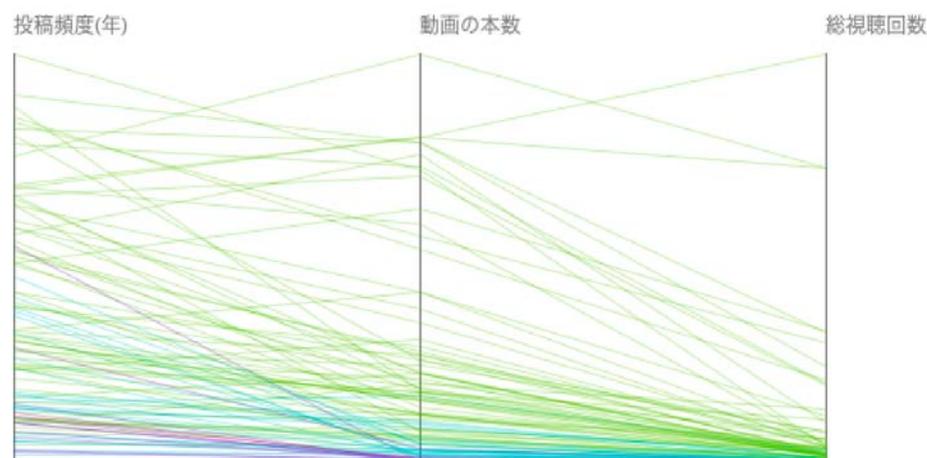
○Clustering=4

わかりやすくするために、
Clustering=4にした。

黄緑、水色、紫、赤の線がある。

「動画の本数」と「投稿頻度」の軸
に注目すると、赤・紫→水色→黄緑
の順に動画の本数が多いことがわか
る。

右上のスライダーを上げる1

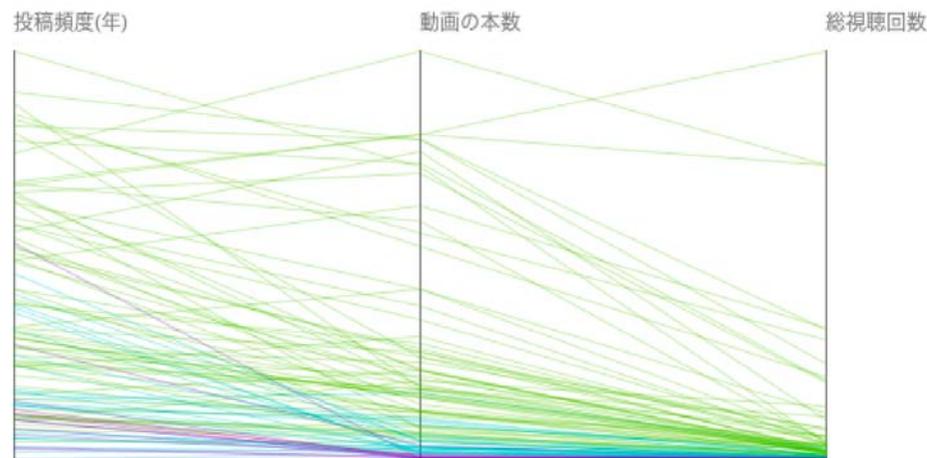


○投稿頻度と動画の本数

投稿頻度が少ないほど、正の相関が強くなり、投稿頻度が多いほど動画の本数にばらつきが出ている（相関が見られない）。

これは投稿頻度を一年あたりに出している動画の本数と定義したため、単にYouTuberを始めた時期による影響を受けたのだと思われる。例えば、最近YouTubeを始めたYouTuber（コンスタントに動画を上げているとする）は、チャンネル開設期間が一年に満たないために、動画の本数はまだ少ないが、投稿頻度は高いという現象が起きてしまう。

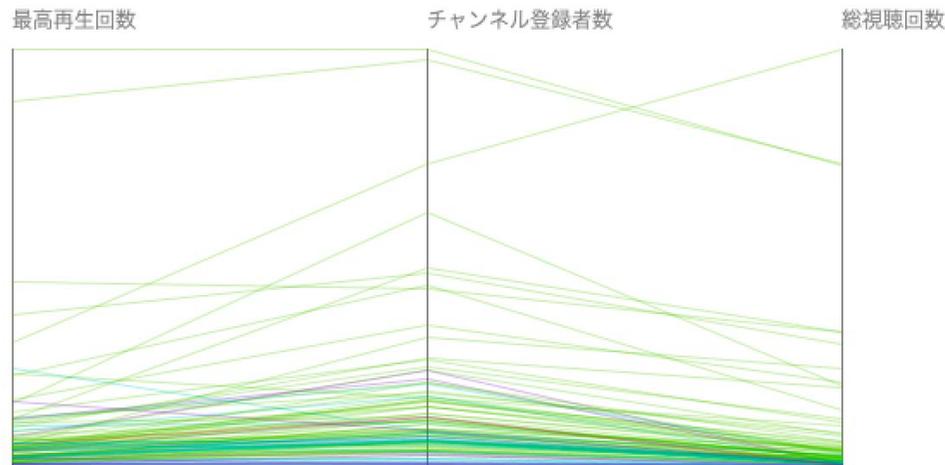
右上のスライダーを上げる1



○動画の本数と総視聴回数

正の相関が見られる。動画の本数が多ければ多いほど総視聴回数が多くなるのは当然のことである。しかし、動画の本数は多くないが、それぞれの動画が爆発的な再生回数を持っている、というタイプのYouTuberはいないことは意外であった。質より量であることがわかった。

右上のスライダーを上げる1

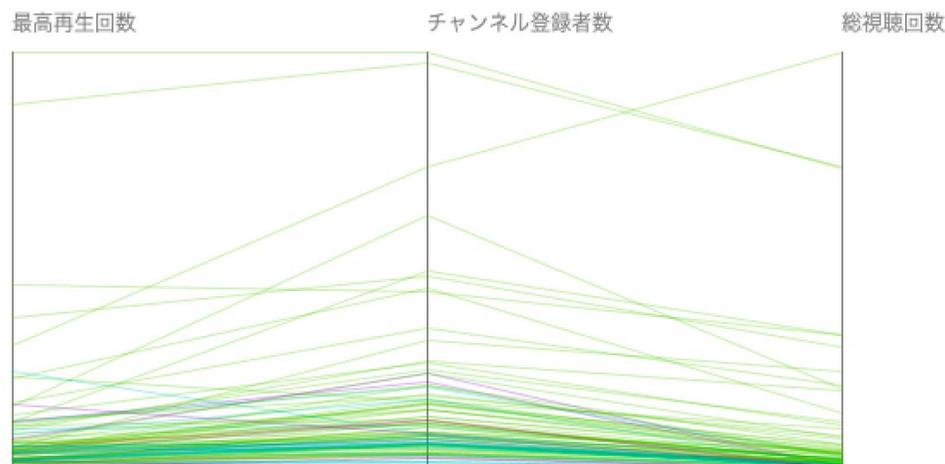


○最高再生回数とチャンネル登録者数
強い正の相関が見られる。

チャンネル登録者数が少ないYouTuberの最高視聴回数が、チャンネル登録者数をはるかに多いYouTuberの最高視聴回数に勝っている線が数本ある。これは「バズった動画を持っているYouTuber」である。動画がバズったからといって、必ずしもチャンネル登録者数が増えるわけではないことが分かる。

紫の線に注目すると、動画の本数が少なくてもチャンネル登録者が比較的多いYouTuberも一部いることが分かる。

右上のスライダーを上げる1



○チャンネル登録者数と総視聴回数

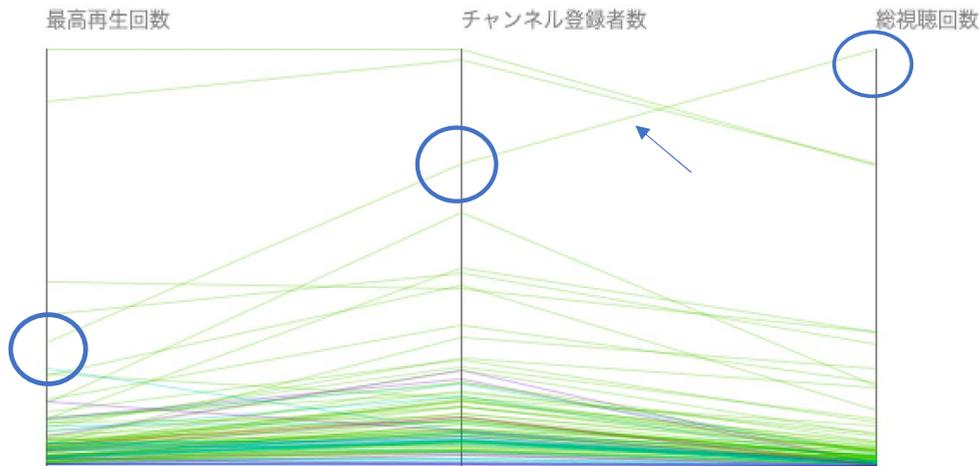
強い正の相関が見られる。

チャンネル登録者数も総視聴回数も多い
YouTuberは動画の本数も多いことが分かる。

チャンネル登録者数が多く、さらに動画の
本数も多ければ、総視聴回数は多くなる。

YouTuberは「**いかにチャンネル登録者数を
増やすか**」が大事だとわかった。

右上のスライダーを上げる1



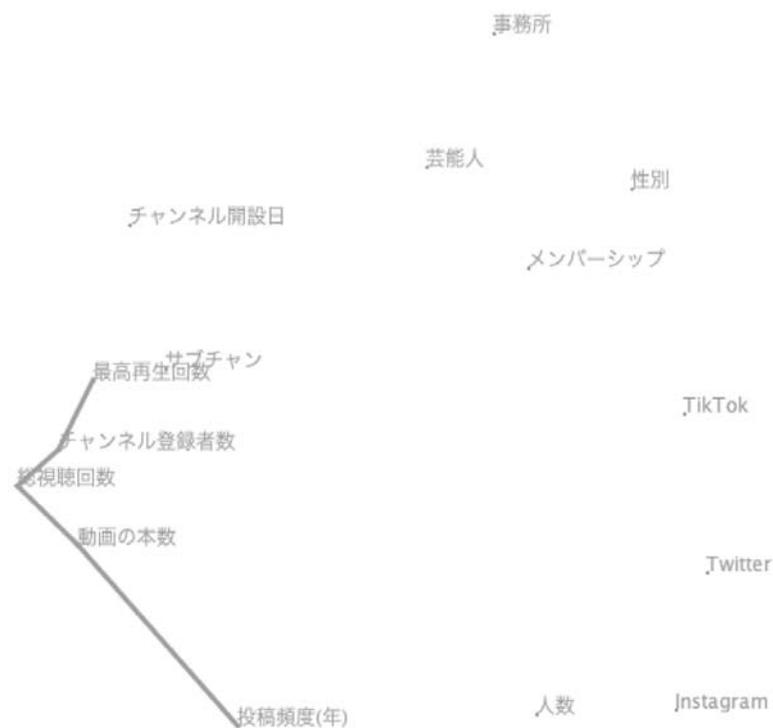
○矢印の線に注目する

このYouTuberはチャンネル登録者数が多い割に最高視聴回数が少ない。しかし、総視聴回数はどのYouTuberよりも多い。

これはどの動画も特別バズるわけではないが、そこそこの再生回数を持った動画をたくさん持っている（線が黄緑である）ということである。

知名度があれば、質より量で勝負することもできると言える。

右上のスライダーを上げる1



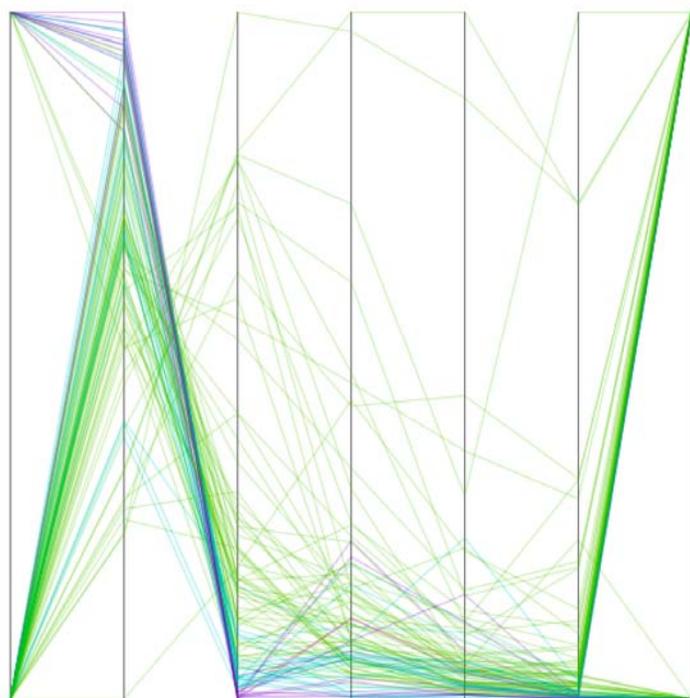
○散布図

チャンネル登録者数と総視聴回数、チャンネル登録者数と最高再生回数、動画の本数と総視聴回数は強い相関があることが分かる。

投稿頻度とチャンネル登録者数の点が離れていることは予想と違い、意外であった。

右上のスライダーを上げる2

芸能人 チャンネル開設動画の本数 チャンネル登録/再生回数 総視聴回数 サブチャン

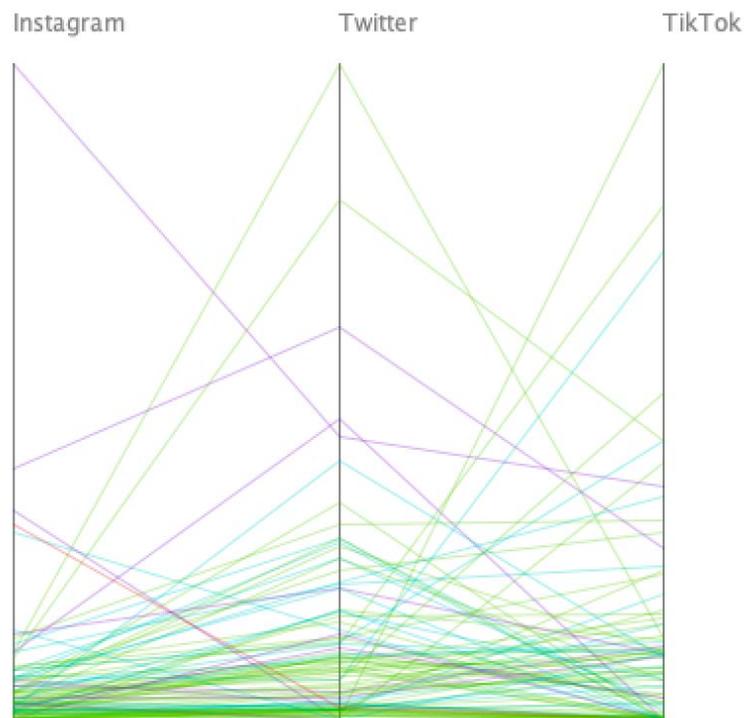


○芸能人とチャンネル開設日

ほとんどの芸能人はチャンネル開設が遅い。

YouTubeに芸能人が参入したのは最近のことであるのが伺える。

右上のスライダーを上げる3



○InstagramとTwitterとTikTok

全体的にフォロワー数が少ない部分は正の相関が強い。

フォロワー数が多くなるほど、相関がなくなる。これは、Twitterは文字媒体を基本としているがTikTokは動画媒体、Instagramは写真媒体を基本としているという違いによるものと思われる。

まとめ

- 動画の本数と総視聴回数は比例の関係にあり、動画の量が要求されていることが分かる
- チャンネル登録者数と最高視聴回数には正の関係がある
- チャンネル登録者数と総視聴回数には正の相関がある
- チャンネル登録者数と総視聴回数が多いYouTuberは動画の本数が多い

予想と反した部分・考察

- サブチャンの有無

YouTuberがサブチャンネルを持つのは、メインのチャンネルである程度のチャンネル登録者数と総視聴回数からだと思っていたため、チャンネル登録者数や総視聴回数と相関があると予想していた。

- メンバーシップの有無

メンバーシップは月額制であるため、メンバーシップの有無は固定ファンが多いかどうかに関係し、チャンネル登録者数や総視聴回数と相関があると予想していた。

予想と反した部分・考察

- 芸能人かどうか

芸能人は元々の知名度があるため、チャンネル登録者数が多いと予想していたが、相関はなかった。知名度があっても視聴者がチャンネル登録をするかは別であるからだと考えた。

- Instagramのフォロワー数

Instagramのユーザーは約3300万人おり、Instagramでのフォロワー数はチャンネル登録者数に影響すると予想していたが、相関はなかった。

InstagramとYouTubeの主要な年齢層が少しずれていることが理由として考えられる。また、女性YouTuberはインスタグラマーを兼ねていることが多いが、男性YouTuberは必ずしもそうではないことも理由に挙げられると思う。

予想と反した部分・考察

- Twitterのフォロワー数

YouTuberユーザーの年齢層とかぶっているため、Twitterのフォロワー数はチャンネル登録者数に影響すると予想していたが、相関はなかったTwitterは文字媒体であるのに対し、YouTubeは動画媒体である。Twitterでバズることが得意な人が動画媒体でもバズることが得意なわけではない。この違いによるものだと考えた。

- TikTokのフォロワー数

YouTubeユーザーと年齢層は違うが、同じ動画媒体であるため、チャンネル登録者数に影響すると予想していたが、相関はなかった。ユーザー層の違いが大きな原因であると思う。さらに、TikTokは動画の長さが最大で1分30秒であるため、動画の長さが10分を超えるのが望ましいYouTubeとはコンテンツに差があると考えた。

予想と反した部分・考察

Instagram、Twitter、TikTokの相関を見た際に(ページ33)、3つのSNS全てで多くのフォロワー数を所持するYouTuberがないことがわかった。

どのSNSでもフォロワー数が多いことはインフルエンサーとしての指標になることは間違いないのだが、その指標はSNS間で共有できるものでないと言える。

よって媒体の違いがYouTuberとの相関の無さの原因であったと考えられる。

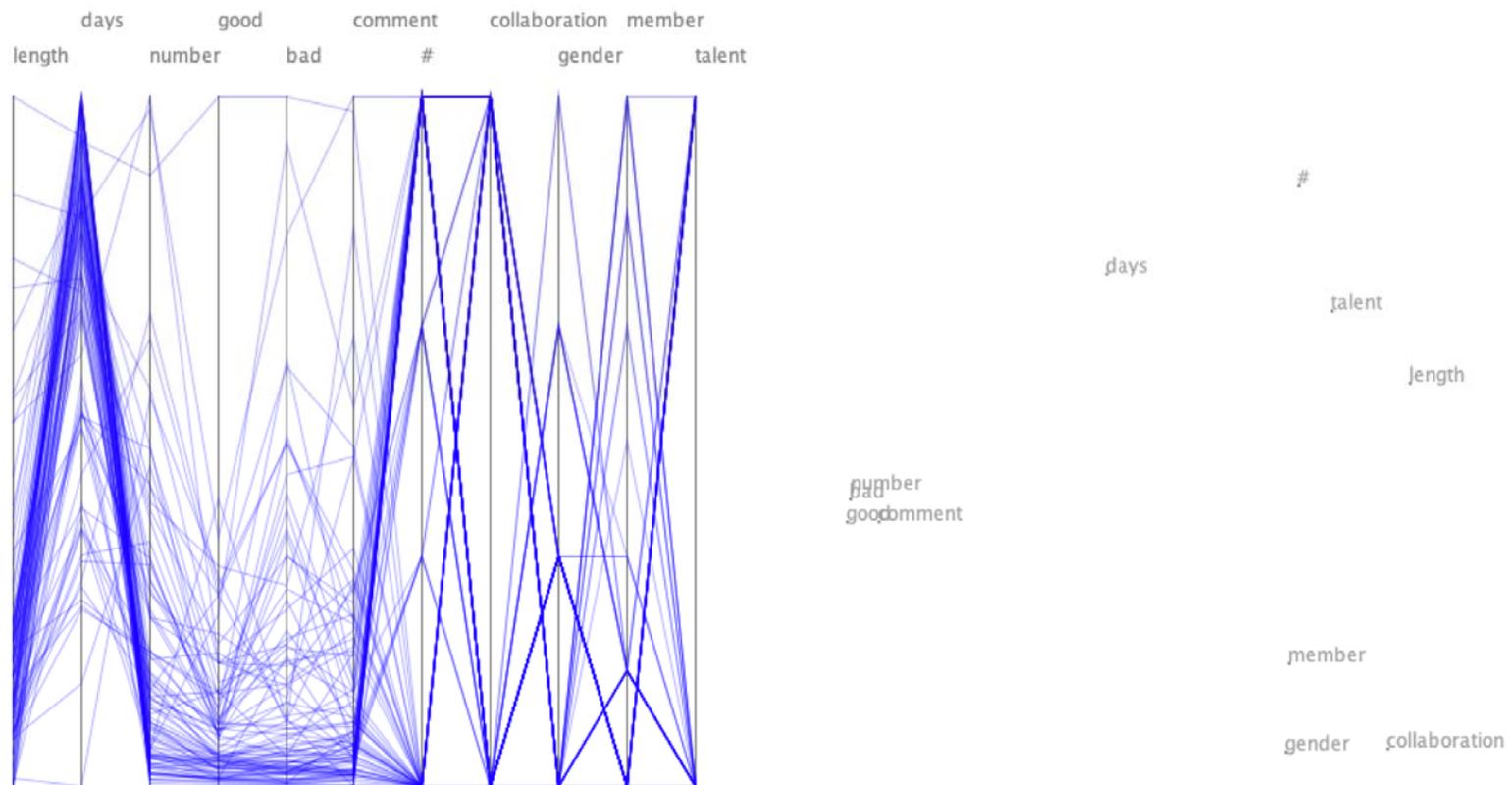
結論

YouTuberとして活躍するには…

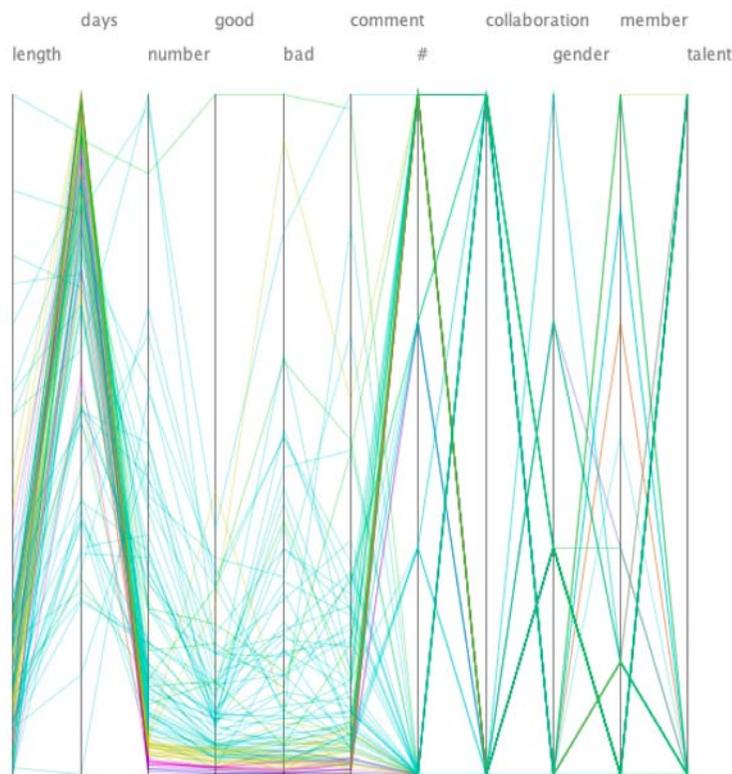
- チャンネル登録者数と動画の本数を増やせば、総視聴回数は必然的に多くなる。
- したがって、チャンネル登録者を増やすことができれば、YouTuberとして活躍することにつながる
- チャンネル登録者数を増やすには以下の方法があるとわかった
 - 動画をバズらせてきっかけを作る
 - ある程度の再生数をコンスタントに維持し、動画の本数を増やす

YouTubeに関する可視化

実行結果



クラスタリング



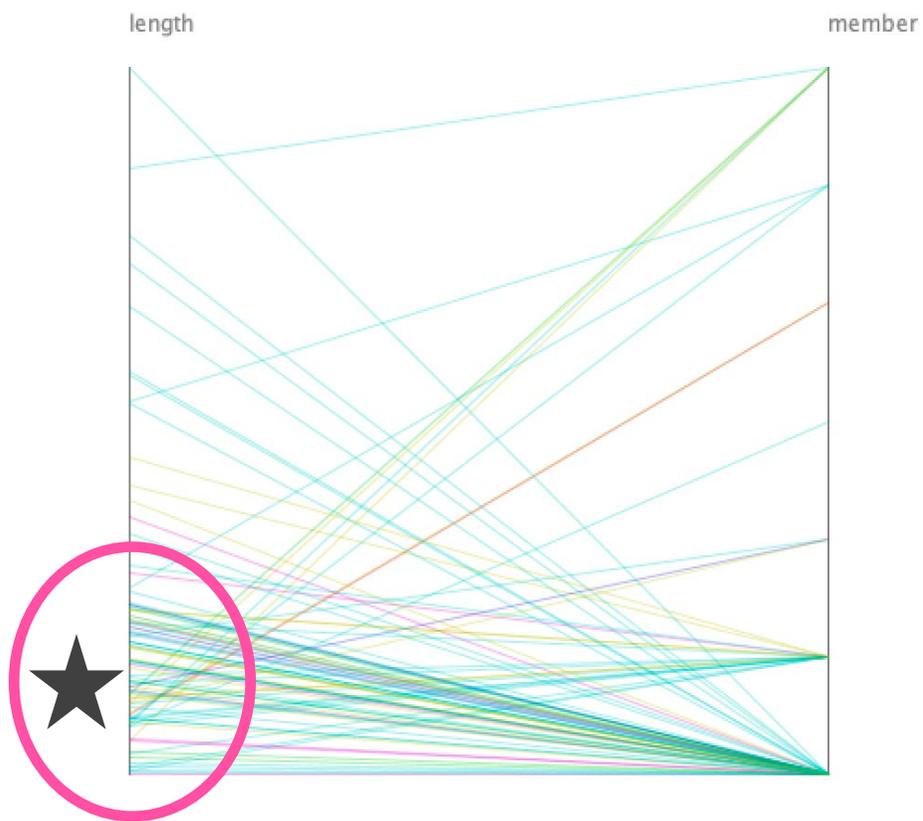
○Clustering=4

わかりやすくするために、
Clustering=6にした。

水色、黄緑、黄色、ピンク、紫、赤
の線がある。

「再生回数(number)」に注目すると、
赤→紫→ピンク→黄色→黄緑→水色
の順に動画が多いことが分かる。

Culling1(右から左にスライダーを移動させていく)



○動画の長さとはYouTuberの人数

特に相関は見られないが、人数に注目して
見てみる。

1人で活動しているYouTuberは、動画の長さはまちまちである。グループで活動しているYouTuberよりコンテンツが絞られるため、動画の長さは短くなる傾向にあると思っていたため、意外であった。長い動画は水色の線なので視聴回数も多い。

大半の動画は人数に関係なく同じくらいの長さになっていることが分かる(★)。YouTubeに適した動画の長さがあることが分かる。

適切な動画の長さ

2020年1月にYouTubeが公式チャンネルで最適な動画の長さについて公開している。内容をまとめると、

- YouTubeの検索結果で上位に来るのは、視聴者が満足できる動画であるかどうか1番の基準である
- 満足度を測るのはどのくらい長く視聴されたかである
- 長い動画は満足度が高く、YouTubeのアルゴリズム上有利になる
- 動画が無駄に長くすることはパフォーマンスが下がるため良くない
- 視聴者を満足させられるなら、動画は長い方がいい

(引用：<https://youtu.be/G1tHzUGdMwY>)

(参考：<https://www.iscle.com/web-it/g-drive/youtube/movie-10minutes.html>)

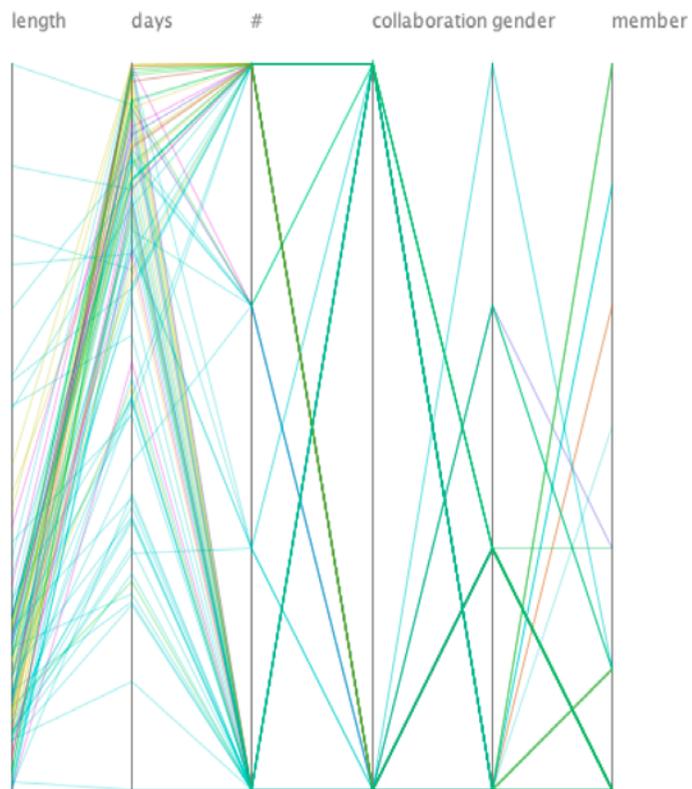
適切な動画の長さ

YouTubeは、**動画の長さはコンテンツに合わせるのがベスト**だと言っている。

多くのYouTuberは、広告収入に影響があるため10分以上の動画にし、さらに視聴者側が満足するような、長すぎず短すぎずの動画の長さにまとめていると考えられる。

動画の全体ではなく、一定時間視聴されれば再生回数に含まれるため、動画の長さが長くても視聴者が見るような満足度のある動画は、動画の長さが長くなっているということである。

Culling2(右から左にスライダーを移動させていく)



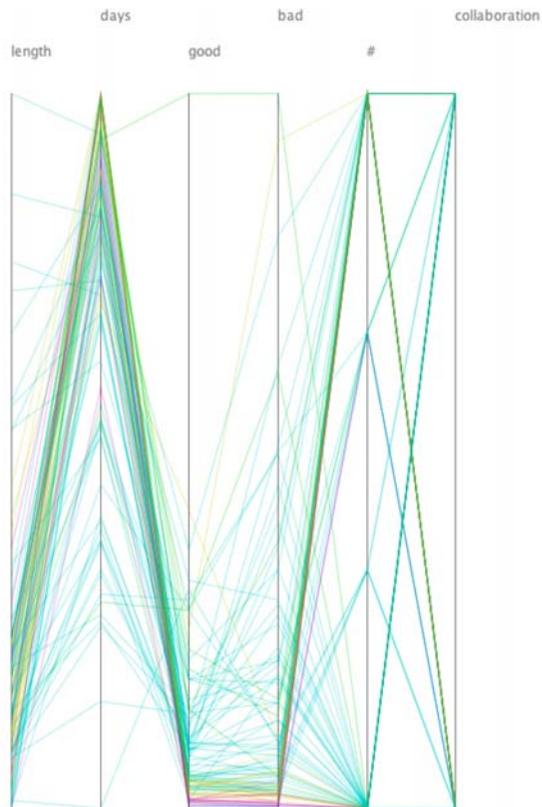
○動画の投稿日とハッシュタグ

特に相関は見られないが、ハッシュタグの個数に注目してみる。

ハッシュタグが一つでも付いている動画は最近の動画が多い。ハッシュタグをつけることがポピュラーになったのは最近であることがわかる。

ハッシュタグが多い方が検索に引っかかりやすくなるため再生回数が多くなると思っていたが、最近投稿された動画の中でハッシュタグの個数と線の色を比較してみると、ハッシュタグの数は再生回数にあまり影響しないことが分かる。

Culling3(右から左にスライダーを移動させていく)



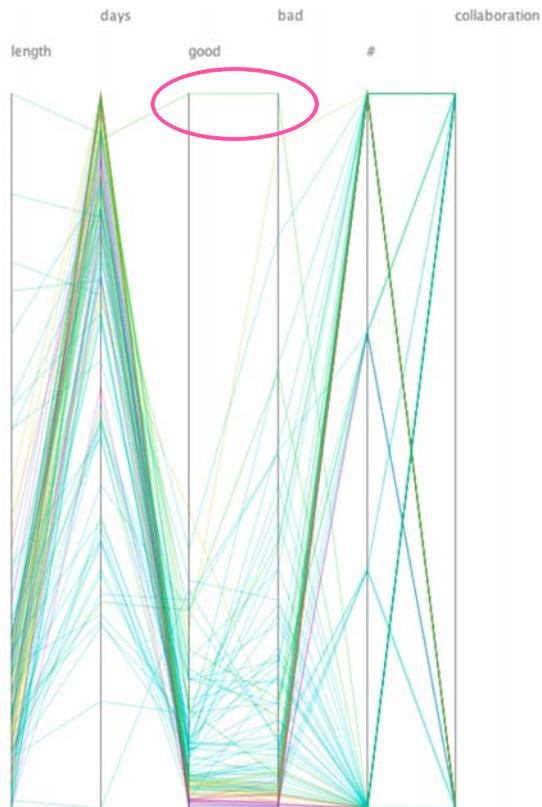
○高評価数と低評価数

再生回数が多い動画は弱い負の相関が見られる。一方で再生回数が少ない動画は正の相関が見られる。

高評価数と低評価数は反比例の関係になるはずである。

高評価と低評価はYouTubeに登録をしていない視聴者でも1クリックで簡単に評価することができるため、信憑性は低い。しかし再生回数が多くなればなるほど信憑性が増す指標であることがわかった。

Culling3(右から左にスライダーを移動させていく)



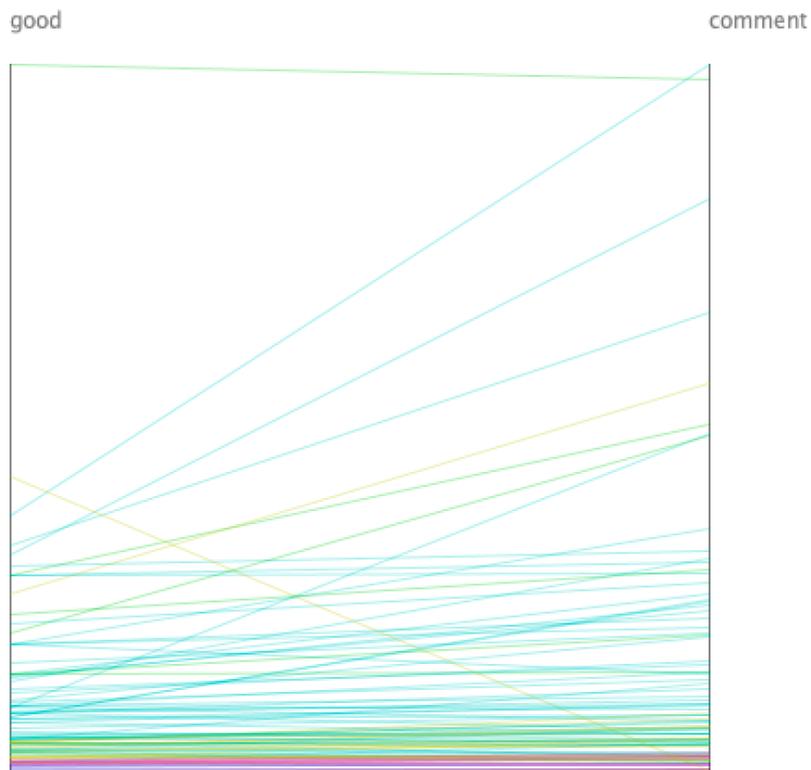
高評価数・低評価数がともに他の動画に比べてかなり多い。

低評価数だけが多い動画は炎上した動画である可能性が高いが、高評価数も多いことから、炎上動画ではないと思われる。

観る人によって意見が割れる動画であったと思われる。原因は以下のことが考えられる。

- ・ 視聴者の年齢層
- ・ 視聴者が動画を見た時期
- ・ 視聴者がアップしたYouTuberのファンかどうか

右上のスライダーを上げる1

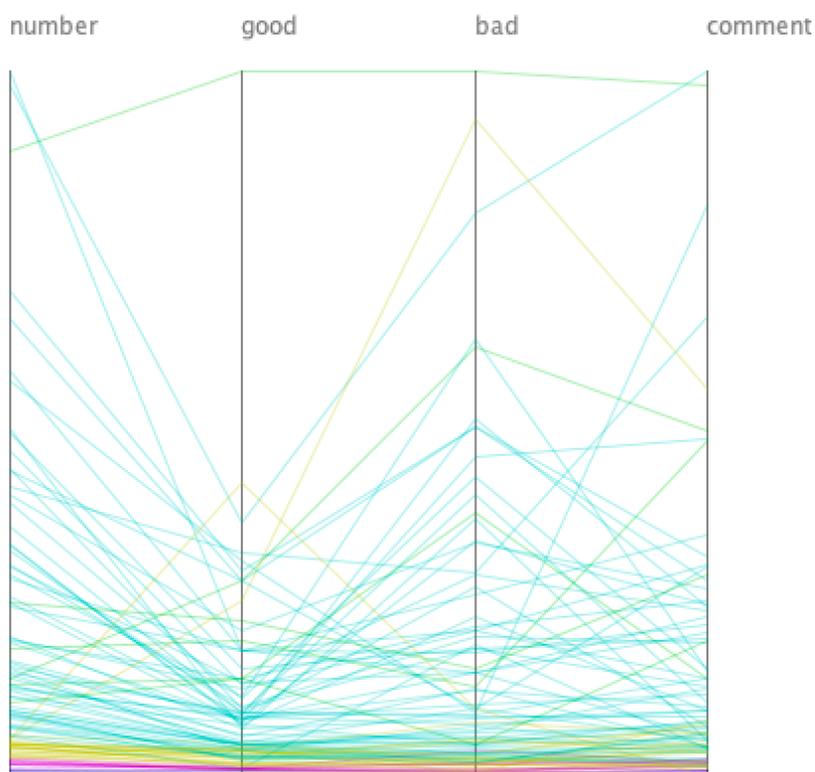


○高評価数とコメント数

全体的に弱い正の相関である。再生回数が少なくなればなるほど正の相関が強くなる。

再生回数が多くなると高評価数とコメント数のどちらも増えることが分かる。

右上のスライダーを上げる2

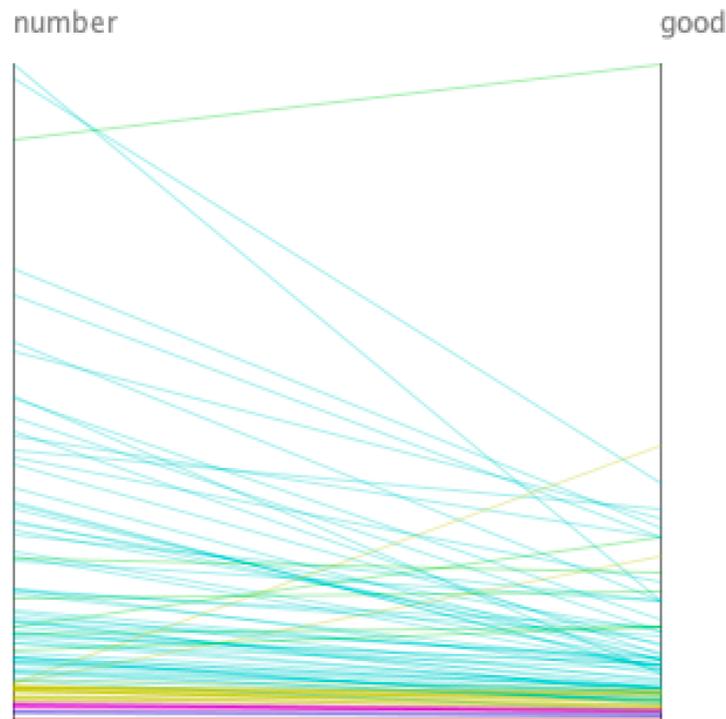


○低評価数とコメント数

高評価数とコメント数の時は弱い正の相関が見られたが、低評価数とコメント数は相関が見られない。

低評価数が多い動画は炎上動画の可能性が高いため、コメント欄が荒れると思っていた。正の相関にならなかったのは意外である。

右上のスライダーを上げる2



○再生回数と高評価数

全体的に弱い正の相関がある。再生回数が少なくなればなるほど、正の相関が強くなる。

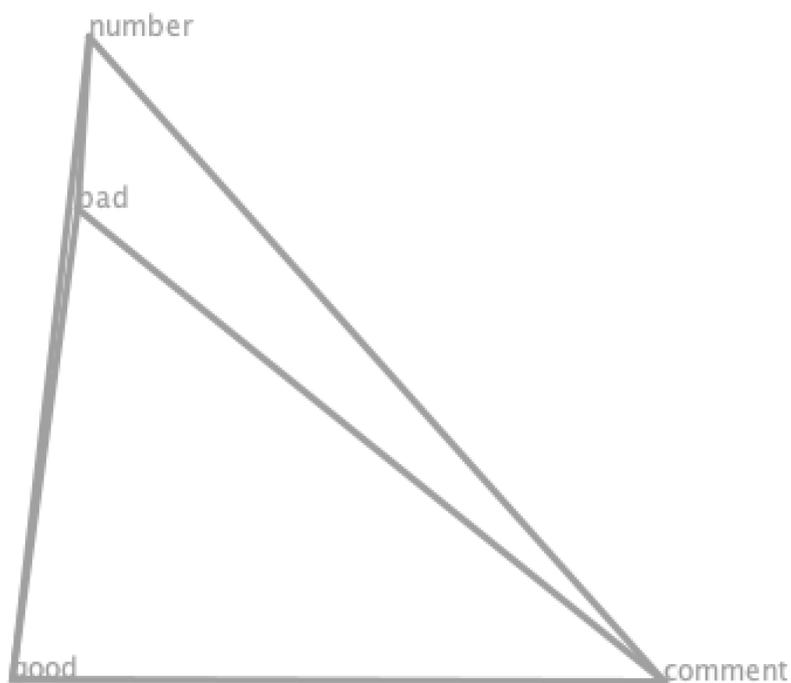
これは、再生回数が多い動画の方が、ファンだけでなく不特定多数の人が見るため、高評価を押し可能性のある人の層が再生回数に対して少なくなるためだと考えられる。

散布図



再生回数、高評価数、低評価数、コメント数は相関が強いことが分かる。

散布図(拡大)



わかりづらいので、拡大して見てみる。

コメント数と最も点の位置に近いのは高評価数である。

高評価数と低評価数では、低評価数の方が再生回数と点の位置に近い。

高評価数とコメント数の再生回数との点の距離が、低評価数と差が出たのは意外であった。

右上のスライダーを上げる3



○動画の投稿日と再生回数

動画の投稿日が最近に近い部分は、強い正の相関が表れている。

アップされている期間が長ければ長いほど、視聴者が再生する期間が長くなるので、当然の結果である。

ばらつきのあるデータ部分に注目すると、動画の投稿日が最近のものでも再生回数が多いものがある（ピンク矢印）。

右上のスライダーを上げる3

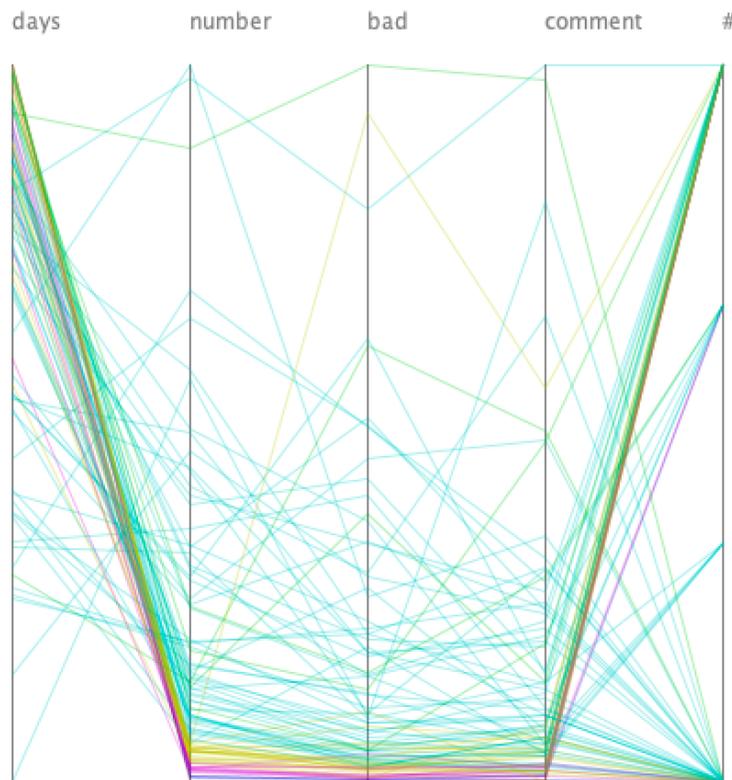


○動画の投稿日と再生回数

ばらつきのあるデータ部分に注目すると、動画の投稿日が最近のものでも再生回数が多いものがある（ピンク矢印）。投稿してすぐバズった動画だと考えられる。

動画の投稿日が昔でも、再生回数が伸びていない動画もある（オレンジ矢印）。再生回数は永久的に伸び続けるのではなく、収束していくことが分かる。

右上のスライダーを上げる4



○再生回数と低評価数

再生回数が少ない部分は正の相関であると言える。

再生回数が多い部分をだんだん見ていくと、再生回数に対して明らかに低評価数が少ない線(…1)と、再生回数に対して低評価数が明らかに多い線(…2)が現れる。

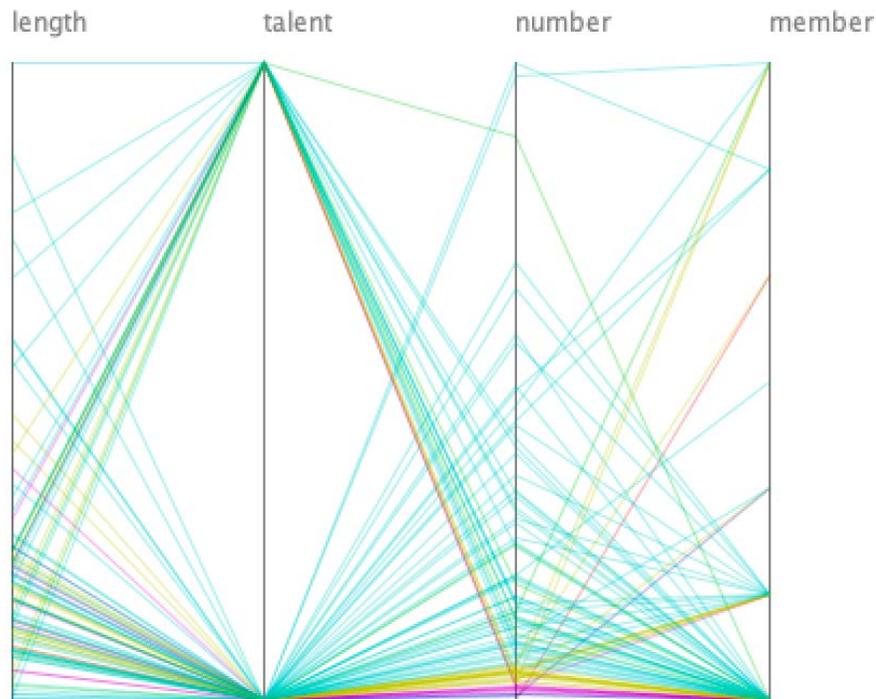
1は単に再生回数が伸びて、かつ、低評価数が少ない動画、と考えていいだろう。

しかし、2に関しては、低評価荒らしが行われた動画である可能性がある。

低評価荒らしとは

- YouTubeの動画を視聴せずに低評価を押す、**組織的**な行為である。これにより、YouTubeは一定時間視聴しないと視聴回数にカウントされないため、視聴回数は増えないが、低評価数は増えるということが起こる。
- YouTubeのクリエイターチームが運営する「Creator Insider」というチャンネルで、この行為を「dislike mobs（低評価する暴徒）」として、何らかの対策を講じる可能性があることを公表した

右上のスライダーを上げる5



○再生回数と芸能人

芸能人は元々の知名度があるため、再生回数が多くなりやすい傾向にあると予想していたが、予想と反して再生回数はバラバラであった。

知名度だけでなく、コンテンツが重要であることが分かる。

YouTubeの視聴者が求めることと、テレビの視聴者が求めていることには違いがあると考えられる。

まとめ

- 動画には適切な長さがあり、大抵のYouTuberはそれを意識している
- ハッシュタグ数は再生回数にあまり影響を与えない
- 再生回数が増えると、高評価数とコメント数のどちらも増える
- 低評価数が多くても、コメント数が増えるわけではない
- 動画の投稿日が早ければ、視聴回数は伸びるが、永久的ではなく収束する
- 芸能人だからと言って再生回数が伸びるわけではない

予想と反した部分・考察

• コラボ動画かどうか

コラボ動画はYouTuber自身の視聴者と、コラボ相手の視聴者の両方が視聴するため再生回数が伸びると思っていた。

コラボはしていても、編集方法や動画のコンテンツはコラボ相手の普段の動画とは違うため、コラボ相手の視聴者を取り込めないからだと考えた。

• YouTuberの性別

女性YouTuberは生活に関する内容（化粧、収納、料理など）が多く、男性YouTuberはお笑いの内容（どっきり、○○してみた、など）が多いイメージがあったため、女性の方がリピート再生が多く、再生回数が伸びると思っていたが、関係しないことがわかった。テレビと違って、動画の長さが短くスキップをすることもできるため、内容に関わらずリピート再生はされているからだと考えた。

結論

再生回数を伸ばすには…

- YouTubeにアップする動画として適切な長さの動画にする
- 高評価数が高くなるような動画にする
- 炎上商法は長い目で見るとやめた方がいい
(YouTube側が検索結果の上位に表示しにくくなる)
- 動画の内容によっては、再生回数は永久的に伸び続けるわけではなく、収束する可能性が高い

感想

YouTubeについて可視化をしてみて、予想通りの結果が出る場合もあれば予想と全く反した結果が出ることもあり、実際に目で見える形にしてデータを見るということの重要さを体感した。

YouTuberは多様他種であるため、データを集める際に、例えば「チャンネル登録者数が10万人以上100万人以下/チャンネル登録者数が100万人以上」のような制限をつけて可視化を行えば、その多様他種さをより詳細に分析できたのではないかと思う。