

コンピュータグラフィックス

Hiddenによる可視化とデータ分析課題

アイドルグループ坂道グループに所属する

メンバーに関する分析

- 概要

今回私は私が日頃応援する日向坂46のメンバーのデータを用いたデータ可視化を行った。しかし、日向坂46は現在32名で、かつ2022年に加入した4期生の12名は表題曲にも参加しておらずデータが不十分である。そのため、同じ坂道グループの乃木坂46、櫻坂46のメンバーのデータも用いて、彼女たちに関する分析を行っていく。

- データについて

今回、2022年12月18日の卒業挨拶を持って活動を終了した私の推しメンである宮田愛萌ちゃんのデータを含めたいため、2022年1月1日時点で在籍したメンバーを対象とした。日向坂46(けやき坂46)の1期生と乃木坂46の3期生、櫻坂46(櫻坂46)の1期生は加入日が近く、ほぼ同期であるといえる。日向坂46の4期生と乃木坂46の5期生、櫻坂46の3期生もまた加入日が近く、彼女たちはみな2022年デビューであり、表題曲に参加しているメンバーはいない。そこで、日向坂46：1,2,3期生の22名と乃木坂46：3,4期生の27名と櫻坂46：1,2期生の23名、計72名を本課題での対象メンバーとした。

以下、表1に本課題で使用した14個のパラメータを示す。

表1：使用したパラメータ

名称	内容
Age	年齢 [歳] 2023年2月6日時点での年齢
Height	身長 [cm] 公式プロフィールを参照
Blood type	血液型 公式プロフィールを参照 A型：1 B型：2

	O型：3 AB型：4
Siblings	兄弟姉妹 一人っ子以外の場合： [対象メンバーは何番目か] / [何人兄弟姉妹か]で算出 例 3人姉妹の2番目 → 2/3 一人っ子の場合：2
From	出身地 公式プロフィールを参照 北海道：1 東北：2 関東：3 中部：4 近畿：5 中国：6 四国：7 九州：8
Songs	参加楽曲数 [曲] 2023年2月6日時点での参加楽曲数 CD収録音源を参照
Formation	フォーメーションに関する指標 2023年2月6日時点での発売されている、または発売が決定している表題曲のフォーメーションを基に決定 詳細は次に示す
Days	活動日数 [日] 加入日から2023年2月6日時点の日数 活動を休止していた場合は、その日数を除いた
Photobook	写真集数 [冊] 2023年2月6日時点で発売されている、または発売が決定している写真集の数 グループ写真集の数+ソロ写真集の数 で算出
Model	専属雑誌数 [冊] 2023年2月6日時点で専属モデルとして起用されている雑誌の数
Followers	SNSフォロワー数 [万人] 2023年2月6日時点でのSNS個人アカウントのフォロワー数 通常の個人アカウントのフォロワー数+写真集アカウントのフォロワー数 (坂道グループのメンバーがソロ写真集を発売する時、その情報を発信する専用のTwitterアカウントが作成されます) ここで、グループのアカウントはこの数値に含めないこととする
Speed	50m走タイム [秒] 彼女たちはレギュラー番組で測定させられるのでそれを参照
Grad	卒業済みか否か 卒業：1 現役：0

Blog	ブログの更新数[回] 2023年1月のブログ更新回数を参照 ただし卒業に伴いすでにブログがクローズされている場合は0とする
-------------	---

フォーメーションに関する指標である**Formation**について説明する。基本的には、CDのタイトル曲のフォーメーションを基に決定する。下図のように各ポジションに番号を振り分けた。

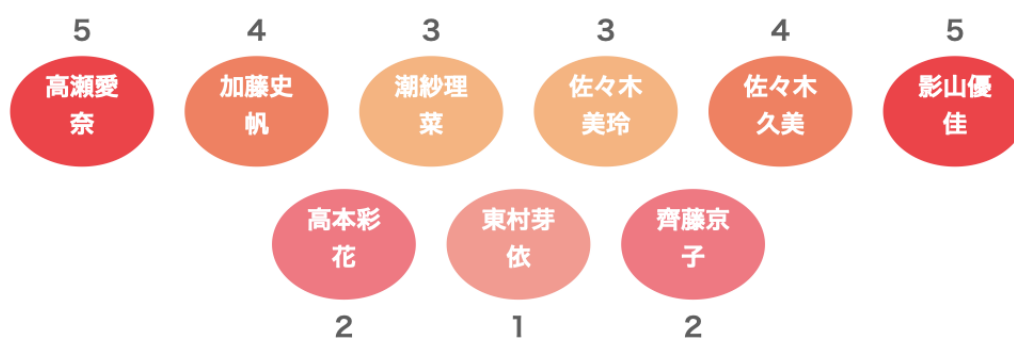


図 1：フォーメーションの例

この図1は日向坂46, 1期生楽曲「どうする？どうする？どうする？」のフォーメーションである。センターは私の推しメンである東村芽依ちゃん。センターは歌割りが多く、MVでのカットインの時間も長い。センター横(高本彩花ちゃんと齊藤京子ちゃん)をフロントというが、フロントはセンターの次に歌割りやカットインが多い。すなわちフォーメーションというのは一般的に、前であるほど、中央であるほど良いとされる。また、左右対称のポジションをシンメという。図の場合、高本彩花ちゃんと齊藤京子ちゃんはシンメである。

Formationは各ポジションに割り振られた番号を、その楽曲参加人数で割ったもので算出する。番号の割り振りは以下に従う。

- センターは1とする
- 一列目から後列へ、中央から端へと番号を振る
- シンメは同じ番号

例えば加藤史帆ちゃんだと、 $4/9=0.4444\dots$ がFormationの値となる。Formationは値が小さい方がより前列で活躍していることになることに注意する。

ここで、坂道グループに限らず、大人数のアイドルグループには選抜制度がある。CDのタイトル曲に参加できるメンバーを選抜メンバーと呼ぶ。現在乃木坂46と櫻坂46は選抜制度を採用している。

乃木坂46では選抜メンバーに選ばれなかったメンバーをアンダーといい、基本的には一つのシングルに、アンダーの曲が1曲収録される。選抜メンバーはタイトル曲で

Formationを算出し、アンダーメンバーはアンダー楽曲でFormationを算出後、その値を1.5倍することで、選抜とアンダーの値に差をつけることとした。

櫻坂46では櫻エイト制度がある。櫻エイトに選ばれたメンバーはCD収録楽曲(ユニット曲を除く)すべてで1,2列目を担当する。櫻エイトに選ばれなかったメンバーをBACKSという。1,2列目の櫻エイトは固定で、3列目のBACKSのメンバーが楽曲ごとに変わるという仕組みである。まず櫻エイトは、タイトル曲とカップリング曲でのFormationをそれぞれ算出し、タイトル曲：カップリング曲 = 1：1.5 となるように計算する。次にタイトル曲に参加したBACKSは、Formationを算出し、そのまま使用する。ユニット曲に参加したBACKSはFormationを算出し、その値を1.5倍した。

このようにして各シングルに対してFormationを計算し、それらの平均が本課題で使用するデータとなる。図2はFormation算出に使用した表のスクリーンショットである。上から順に日向坂46、櫻坂46、乃木坂46である。こうしてみると乃木坂46の活動が最も盛んであることがわかる。

The image displays three screenshots of spreadsheets used for formation calculations. The top screenshot is for 日向坂46, the middle for 櫻坂46, and the bottom for 乃木坂46. Each spreadsheet has multiple columns, including member names, formation numbers, and calculated values. The data is organized in a grid format, with rows representing different formations and columns representing individual members and their respective values.

図2：Formation算出の様子

● 結果・考察

結果は以下の図3の通りである。

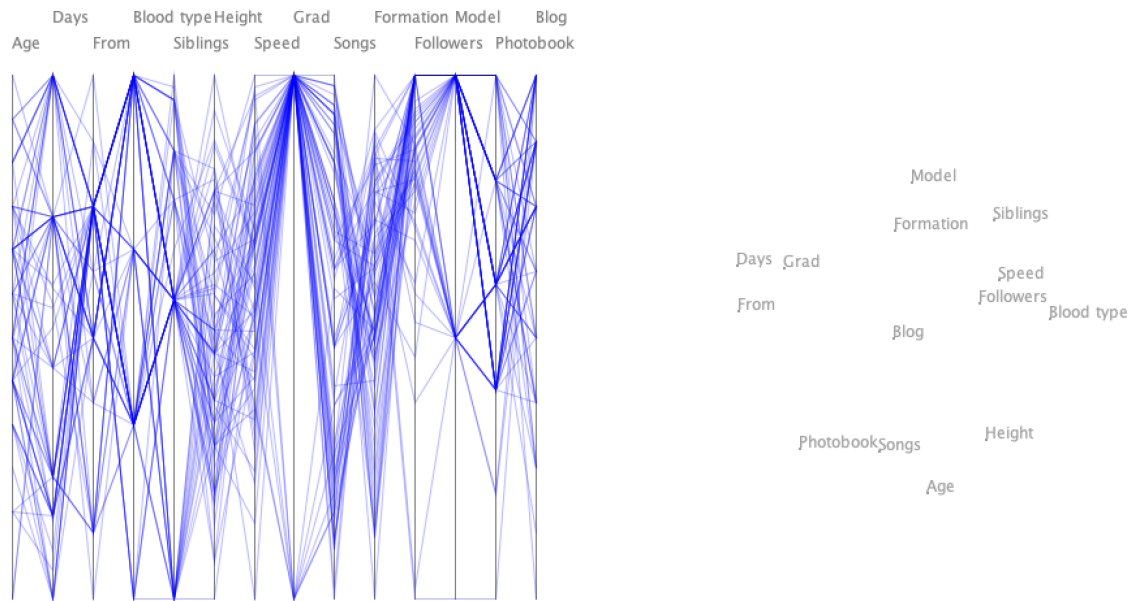


図3：解析結果

散布図に注目すると、**Blog**が中央付近に配置され、どのパラメータにも近いとは言えないように思えた。そこで、右のスライダーを散布図やPCPに変化があるまで動かしたところ、図4のようになった。

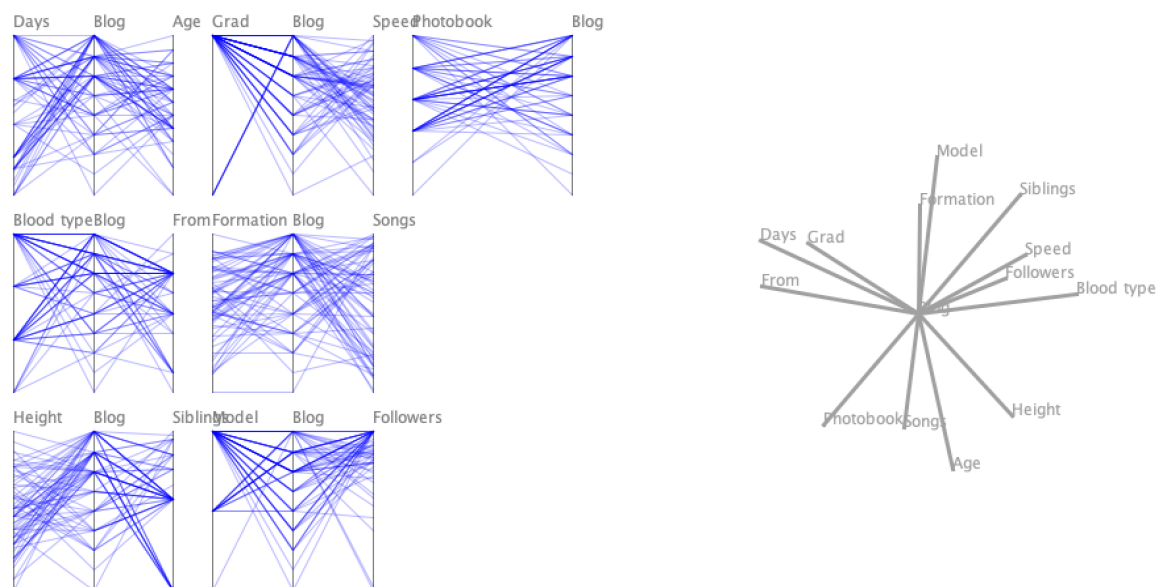


図4：スライダー操作後

散布図を見れば**Blog**がその他すべてのパラメータと接続されているため、**Blog**と各パラメータの相関の強さは近いものであると考えられる。また、PCPについても相関は見られない。よって**Blog**はどのパラメータに対しても無相関であり、無視しても良いのではないかと考えた。

ここで、一度**Blog**を削除したデータを表示してみた。図5に示す。

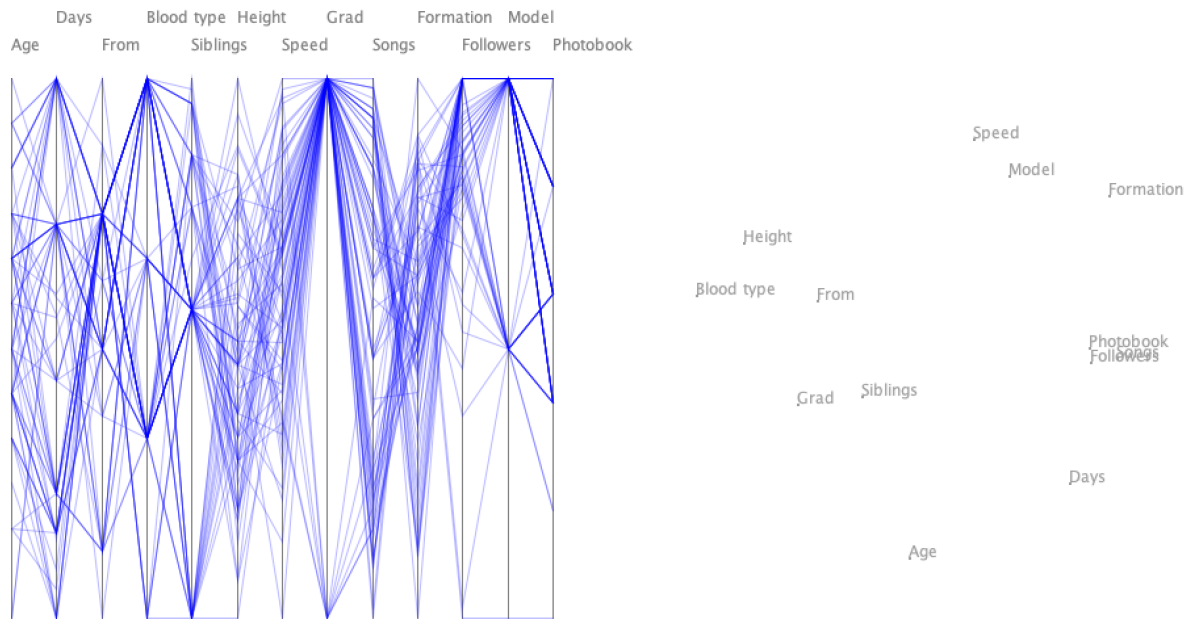


図5 : **Blog**を削除した場合の解析結果

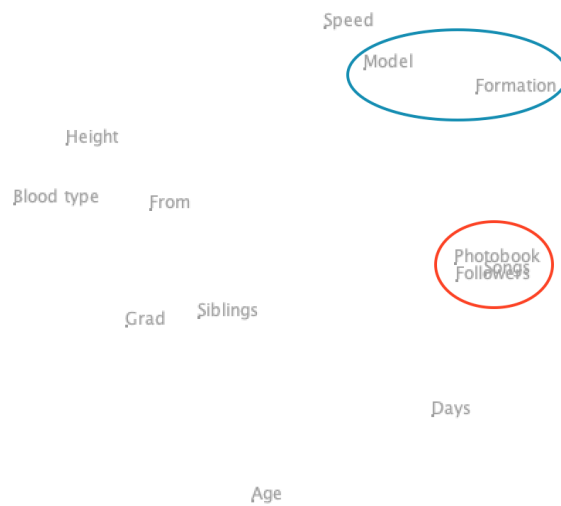


図6 : 散布図

図6に注目すれば、赤色の楕円で囲まれた部分の**Songs**、**Photobook**、**Followers**の相関が高いことがわかる。人気メンバーであるほど、参加楽曲数(**Songs**)は多く、ソロ写真集(**PhotoBook**)を出す可能性も高い。また、ソロ写真集を出せばSNSのフォロワー数(**Followers**)が多くなるため、この結果は容易に想像できる。一方で、これらは青色の楕円で囲まれた部分の**Model**や**Formation**との強い相関関係がみられなかった。**Model**

や**Formation**は、**Songs**、**Photobook**、**Followers**と同様にそのメンバーの人気のそのまま影響するパラメータであると考えられたため、少々意外な結果であった。

次に2つのクラスターにクラスタリングした。その結果を図6に示す。

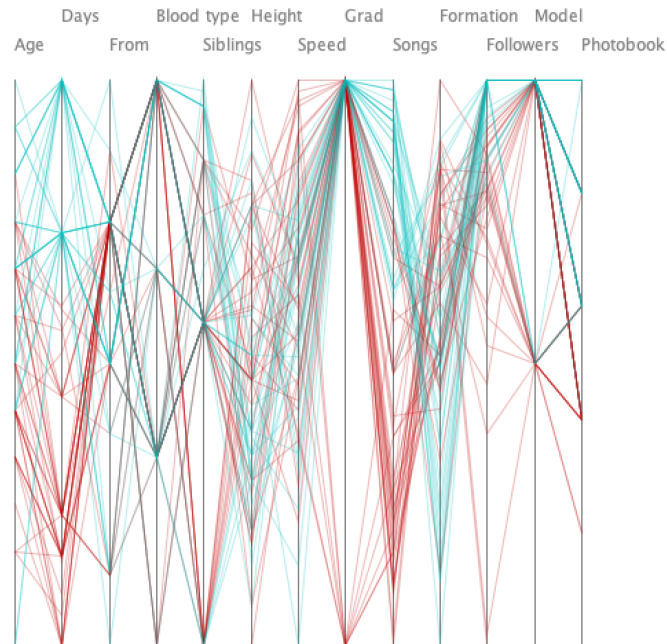


図6 : Clustering 2 の結果

Days、**Songs**、**Photobook**に注目すると値の大小でクラスターが二分されているようにみえる。活動日数(**Days**)の長さは、参加楽曲数(**Songs**)や写真集数(**Photobook**)に影響すると考えられる。これらの関係がクラスタリングすることで明確となったため、クラスタリングしたままデータ分析を行うこととした。赤色の線は**Days**、**Songs**、**Photobook**で大きい値をとっており、青色の線は小さい値をとっているため、ベテラン(赤色)と若手(青色)と大別することにする。ここで、ベテランであるから年齢が高く、若手であるから年齢が低いというわけではないことに注意したい。例えば在籍6年を超える乃木坂46の岩本蓮加ちゃんは12歳で加入したため、グループの中でも若い。

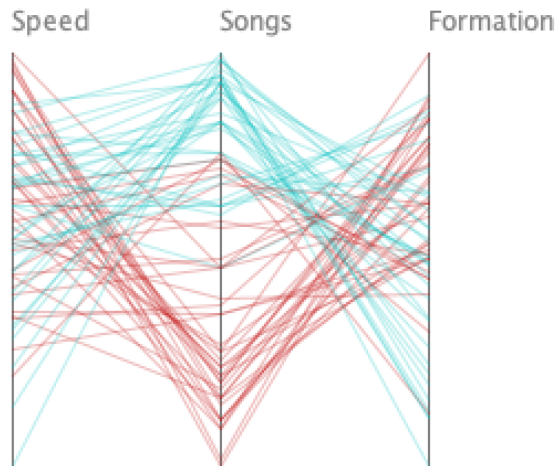


図7 : Speed, Songs, Formationの関係

右上のスライダーを操作したところ、図7のようなPCPが得られた。SpeedとSongsとFormationの関係が示されているが、クラスタを考慮すれば正負の相関が確認できる。

SongsとFormationについて、ベテランには負の相関があり、若手には正の相関がある。これより、ベテランが世代交代の隙を与えず根強く最前線で活躍していることがわかる。また、若手の中でもSongsが高い場合、Formationは小さな値をとっており、新世代がこの結果を形成しているのだろう。

また、SpeedとSongsについて、ベテランには正の相関があり、若手には負の相関がある。どうやらベテランはアイドルとして人気があるだけでなく、走る能力にも長けているらしい。特に日向坂461期生は運動能力お化けが多く在籍しており、日本の成人女性50m走平均タイムである9.04秒[1]を全員が上回る結果を叩き出している。その影響もあるのだろうか。とは言え、アイドルの醍醐味「歌って踊る」というのは体力的にかなり大変なものであることは間違いないし、体系維持のための日々の運動や、ステージ上での全力疾走なども多々見受けられるので、自然とそのような力がついたのかもしれない。

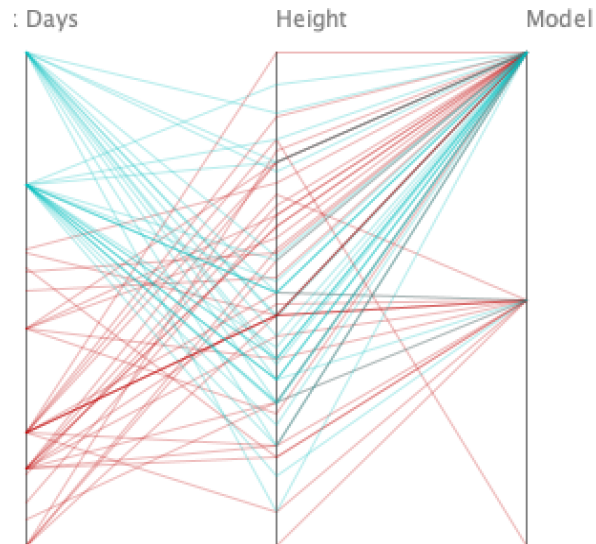


図8 : Days, Height, Modelの関係

次に図8に注目したい。DaysとHeightとModelの関係が示されている。

DaysとHeightを見れば青色に正の相関が、赤色には若干の負の相関がある。若手は比較的身長の高い子が多いということだろう。若手＝若いではないが、対象メンバー内で一回りも年齢が違うこともあり、日本人の平均身長は増加傾向にある[2]。そのため、この傾向が同じグループ内でありのままに現れたのだとも考えることができる。

HeightとModelを見れば、負の相関があるように見える。モデルに選ばれているメンバーは比較的身長が高いメンバーが選ばれているものの、外れ値もみられ、スタイルよりも顔重視の起用である可能性は否めない。(お顔も大切ですよ)

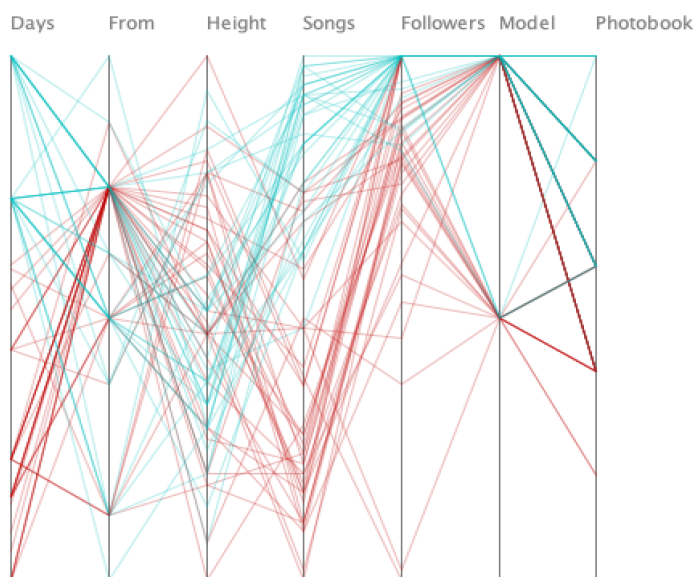


図9 : Days, From, Height, Songs, Followers, Model, Photobookの関係

ここで図9に注目したい。

Heightと**Songs**を見ればベテランは比較的正の相関があり、若手には負の相関がある。若手は活動日数が少ないため、このような結果になることは容易に想像できる。一方で、ベテランは身長が高いほど参加楽曲数が多いことがわかる。**Songs**はグループ内の人気はかなり影響することはすでに述べている通りであり、坂道グループ内での選抜や良いポジションに選ばれやすい要素の中に身長もまた影響することが確認された。

Daysと**From**を見ると、ベテランは関東(3)からの選出が多いことに対して、若手は関東に限らず全国から選出されていることがわかる。これはやはり知名度の拡大から現れるものであろう。かつてはテレビ東京でしか放映されていたなかった、各グループのレギュラー番組がYoutubeやひかりTVで配信されるようになったり、グループ規模の拡大など、様々な要因は考えられるが、ファンとしてはとても嬉しい解析結果である。

● 感想

様々な角度から自身の応援するグループを解析し、グループについてさらに詳細な知識を得て、応援するモチベーションもまた高まりました。

データの収集が困難であることから諦めましたが、メンバーの骨格や、メンバー別有料メッセージの購読者数ランキングなども使用して、データ解析を行いたかったです。これらの結果を組み合わせることで選抜に選ばれるための基準を明確にして、事務所の良くないところを暴いてやろうという野望(?)込みで...

伊藤先生の課題では、必ず自分の好きなコンテンツを取り上げてきたため、課題をしているにも関わらず、とても楽しく充実した時間でした。

● 参考文献

- 日向坂46 OFFICIAL WEB SITE.
(<https://www.hinatazaka46.com/s/official/?ima=0000>)
- Wikipedia. 「日向坂46」.
(<https://ja.wikipedia.org/wiki/%E6%97%A5%E5%90%91%E5%9D%8246>)
- 乃木坂46 OFFICIAL WEB SITE.
(<https://www.nogizaka46.com/s/n46/?ima=1553>)
- Wikipedia. 「乃木坂46」.
(<https://ja.wikipedia.org/wiki/%E4%B9%83%E6%9C%A8%E5%9D%8246>)
- 櫻坂46 OFFICIAL WEB SITE. (<https://sakurazaka46.com/s/s46/?ima=1514>)
- Wikipedia. 「櫻坂46」.
(<https://ja.wikipedia.org/wiki/%E6%AB%BB%E5%9D%8246>)
- 「【日向坂】メンバーの兄弟姉妹を一覧表にしたよ!【家族構成がわかる】」.
2021/7/4. (<https://love-koba.com/entry/hinata-brother-sister>)
- 櫻坂46まとめ速報. 「【櫻坂46】メンバーの兄弟姉妹の一覧」. 2021/5/2
(<http://sakurazaka46matome.com/archives/9597747.html>)
- 乃木坂ラジオの時間. 「【乃木坂46】乃木坂メンバーの兄弟構成を今一度確認」.
2017/11/21. (<http://nogiradi.com/archives/5083958.html>)
- おひさまポカポカ日誌. 「【日向坂46】【櫻坂46】50m走まとめ」. 2020/11/29.
(<https://ameblo.jp/yoogawa14sunriver/entry-12640873128.html>)
- あの坂道をのぼれ. 「乃木坂46メンバーの50m走タイムランキング【2022年46時間TV+5期生体力測定】」. 2022/8/10.
(<https://anosaka.com/nogizaka46-50m-ranking>)

[1] タクトレブログ. 「50メートル走の平均タイムは?年齢や学年/男女別に速い~遅いを9段階評価 - 大人の50メートル走平均タイム-」
(<https://re1wa018.com/50m-average/#:~:text=%E8%A6%8B%E3%82%89%E3%82%8C%E3%81%BE%E3%81%99%E3%80%82-,%E5%A4%A7%E4%BA%BA%E3%81%AE50m%E8%B5%B0%E5%B9%B3%E5%9D%87%E3%82%BF%E3%82%A4%E3%83%A0,%E3%82%BF%E3%82%A4%E3%83%A0%E3%81%AF9.04%E7%A7%92%E3%81%A7%E3%81%99%E3%80%82>)

[2] 社会実情データ図録.「日本人の平均身長・平均体重の推移」.
(<http://honkawa2.sakura.ne.jp/2182.html>)