

情報可視化ソフトウェア Hidden による 韓流アイドル チョン・ウォヌのチッケム についてのデータ分析

1. 概要

私は韓流アイドルグループ「SEVENTEEN」のメンバーである、チョン・ウォヌ氏(以下ウォヌ)のファンである。日頃から彼の動画を拝見しているが、中でも”チッケム[*1]”と呼ばれる映像は彼一人だけにカメラを向け、踊ったり歌ったりする様子を高画質で見られるため、頻繁に楽しませてもらっている。そこで、ウォヌのチッケムについて、どのような要素がその再生数やコメントの数に影響するのか気になり、調べてみることにした。

[*1]…主に 1 つのカメラでアイドルやアーティストのパフォーマンスを撮った映像のことを指す。アイドルグループの場合は特定のメンバーにフォーカスし、その 1 人の姿を追ったものがメジャーである。

COSMOPOLITAN/ [「チッケム 20 選」](#) より/最終閲覧日：2023.1.15

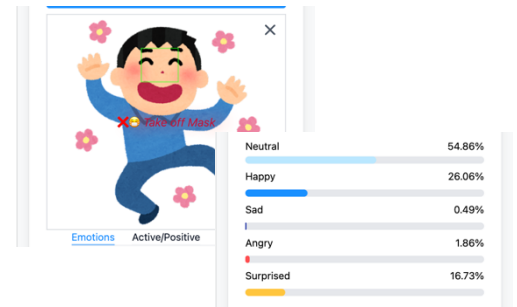
2. データの出典・参考サイト・利用ソフト

YouTube

- 「스브스케이팝 X INKIGAYO」…「Inkigayo」におけるチッケムの収集
- 「KBS Kpop」…「Music Bank」におけるチッケムの収集
- 「STUDIO CHOOM [스튜디오 춤]」…「BE ORIGINAL」におけるチッケムの収集
- 「SEVENTEEN」…「SSSS CAM」におけるチッケムの収集
- 「Mubeat」…「Show Champion」におけるチッケムの収集
- 「MBCkpop」…「Music Core」におけるチッケムの収集
- 「Mnet TV」…「M COUNTDOWN」におけるチッケムの収集
- 「random_k」…チッケム該当曲の Line Distribution の取得

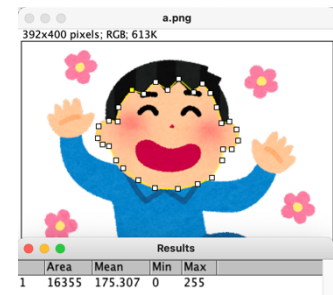
MAL FaceEmotion

- 右図のような形で、サムネイル画像(以下サムネ)のウォヌの表情を数値化



ImageJ

- 右図のような手順で、サムネのウォヌの
”肌の面積”/”頭部の面積”/”体の面積”を計測(手動)
- 髪色の RGB 取得



@shobota/【Python】Youtube の再生リスト ID から動画情報を取得して Excel に書き出す

- YouTube の再生リストより、再生リスト内の動画情報の取得のために参考

<https://il.ytimg.com/vi/【動画 ID】/sddefault.jpg>

- 動画のサムネの取得

3. データ内容

以下の 24 項目を設定。

変数名	内容	単位
Views	各チッケムの再生回数	回
Cmt	各チッケムのコメント数	個
Good	各チッケムの高評価数	個
Days	(各チッケムの公開日)-(該当曲の MV の公開日) [*2]	日
MVV	各チッケムの曲の MV の再生回数 [*3]	回
Nr	サムネのウォヌの表情に含まれる、Neutral の割合	%
Hp	サムネのウォヌの表情に含まれる、Happy の割合	%
Sad	サムネのウォヌの表情に含まれる、Sad の割合	%
Ar	サムネのウォヌの表情に含まれる、Angry の割合	%
Sup	サムネのウォヌの表情に含まれる、Surprised の割合	%
Pos	サムネのウォヌの表情に含まれる、Positive の割合	%

Act	サムネのウォヌの表情に含まれる、Active の割合	%
Vbs	サムネのウォヌの表情に含まれる、Vibes の割合	%
hR	サムネのウォヌの髪色の R 値	bit
hG	サムネのウォヌの髪色の G 値	bit
hB	サムネのウォヌの髪色の B 値	bit
hRGB	サムネのウォヌの髪色の(R+G+B)値 大きい：白に近い/小さい：黒に近い	bit
Dist	各チッケムの曲における、ウォヌのパート配分率	%
DiR	DistRank。チッケム曲において、メンバー中の順位 1：パートが最多/13：パートが最少	位
sA	skinArea。サムネ中のウォヌの肌面積	pixel ²
hA	headArea。サムネ中のウォヌの頭部面積	pixel ²
s/hP	$[(\text{skinArea})/(\text{headArea})]*100$ 頭部面積中の肌面積の割合 大きい：短髪 or セクシーな衣装/小さい：長髪(長め)	%
wA	wonwooArea。サムネ中のウォヌ全体の面積	pixel ²
h/wP	$[(\text{headArea})/(\text{wonwooArea})]*100$ ウォヌ面積中の頭部面積 大きい：接写/小さい：遠写	%

[*2]…アルバムのタイトル曲でない場合、アルバム発売後時間が経ってから MV 発表となる場合があり、負の数にもなりうる。また、カバー曲でのチッケムにおいては直近の MV 公開日からの日数とした。チッケム曲に MV がない場合、収録されているアルバムの発売日からの日数とした。

[*3]…チッケム曲の MV がない場合、“Choreography Video”/“Official Audio”のうち再生回数の多い方を採用した。

4. カテゴリ

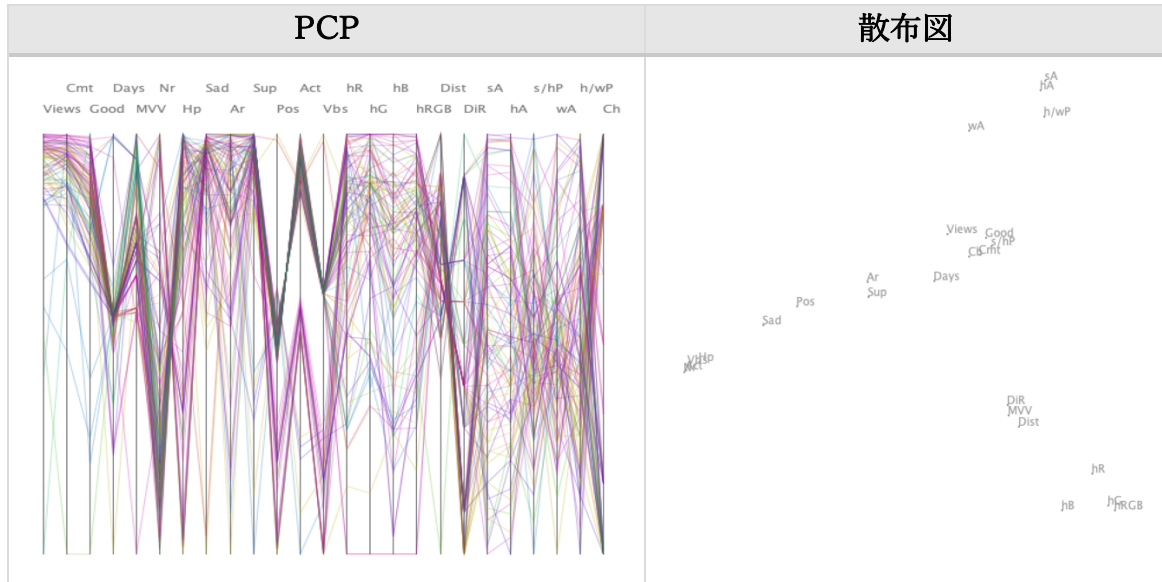
以下のカテゴリを設定。

カテゴリ名	数値	内容
Ch	1. Inkigayo 2. Music Bank 3. BE ORIGINAL 4. SSSS CAM 5. Show Champion 6. Music Core 7. MCOUNTDOWN	チッケム公開チャンネルごとに分類

5. 結果と考察

5.1. 全体

クラスターを7に設定した時のグラフは以下。



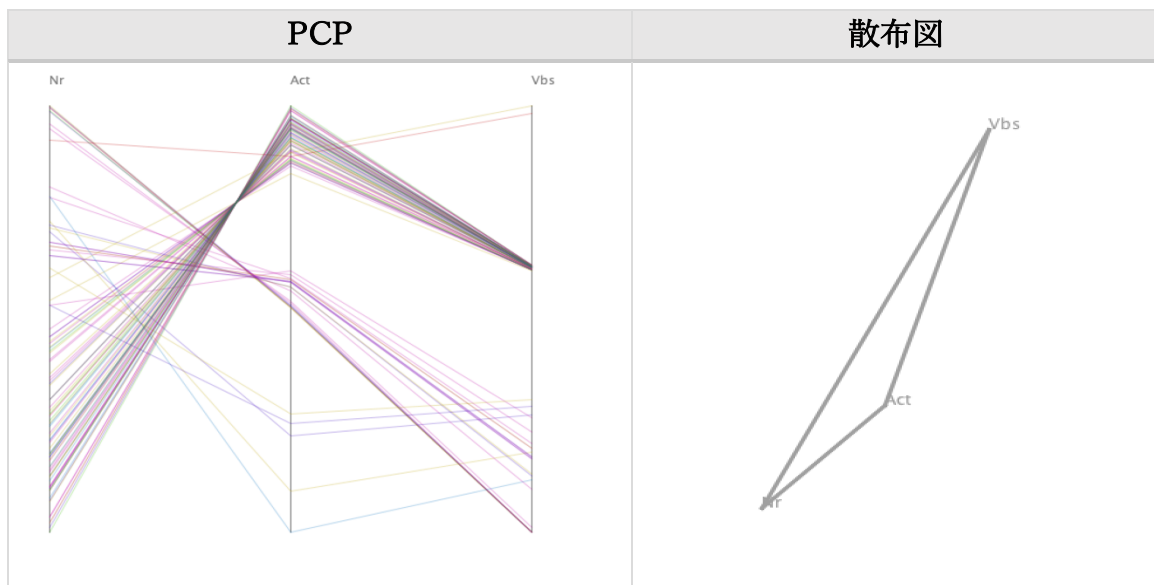
全体を見てもわかりにくいため、右上のスライダーを利用して相関の強いものから見ていく。

強い相関のあるものについて観察したいため、今回は右スライダーのピンクで囲った部分についてのみ、結果を考察していきたいと思う。

相関の強いものから順に、紹介していく。

5.2. 相関の強さ：★★★★★

まず、最も相関の強かったものを見ていこう。



最も相関が強かったものは、[Nr: Neutral][Act: Active][Vbs: Vibes]であった。

<PCP からわかること>

[Neutral]と[Active]には強い負の相関があり、[Active]と[Vibes]には強い正の相関がある。またそこから、[Neutral]と[Vibes]には強い負の相関があることがわかる。

<散布図からわかること>

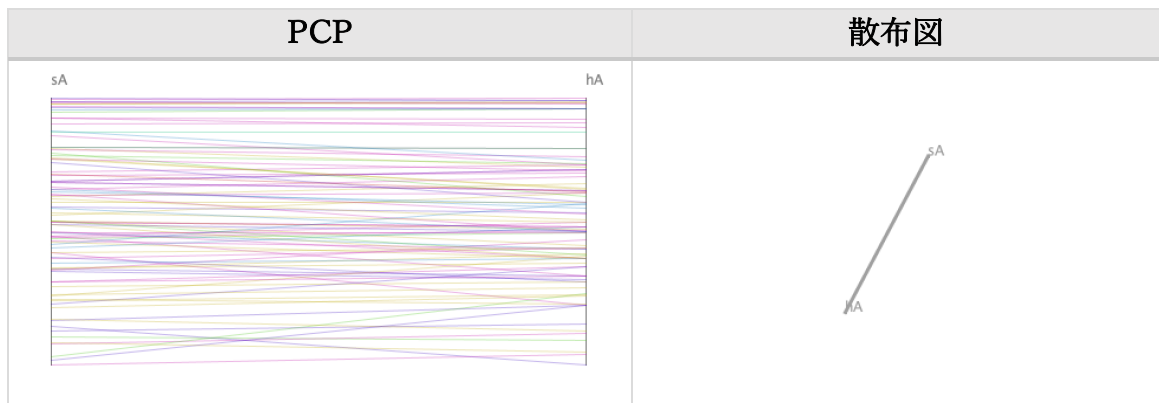
[Neutral], [Active], [Vibes]はそれぞれ互いに強い相関をもち合っている。

<考察>

無表情←[負の相関]→活発な表情/雰囲気がある表情というのは直感的に正しい。

5.3. 相関の強さ：★★★★

次点で、相関が強かったものはなんだろうか。



次に相関が強かったものは、[sA: skinArea][hA: headArea]であった。

<PCP からわかること>

[skinArea]と[headArea]には強い正の相関がある。

<散布図からわかること>

[skinArea]と[headArea]には強い相関がある。

<考察>

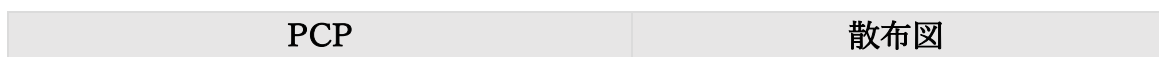
肌の面積←[正の相関]→頭部面積というのは直感的に正しい。

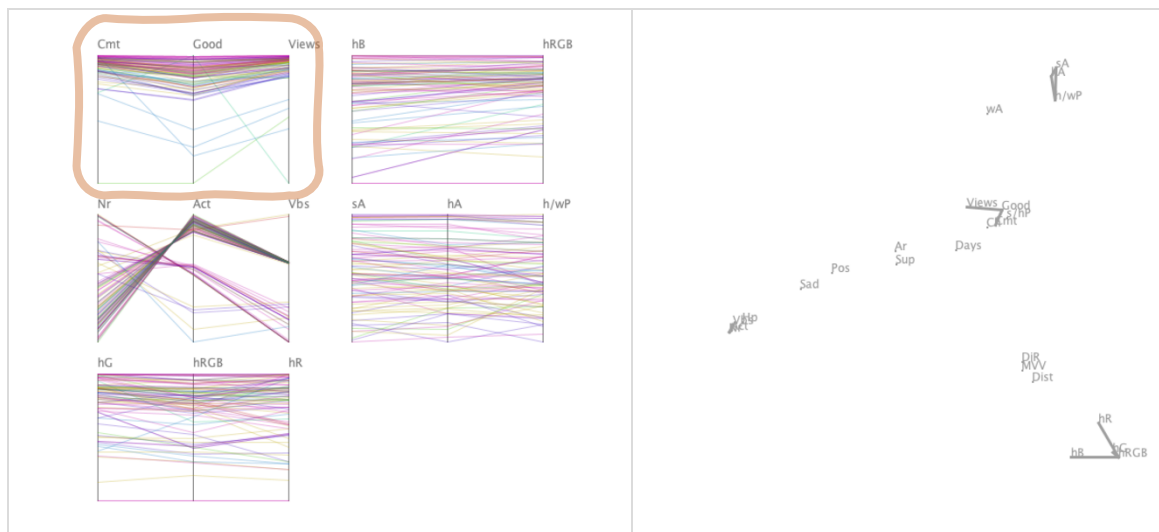
相関の強いものから順に見ていくと、いわば自明とも言えることばかりであった。
しかし、この後から徐々に面白い結果が見えてくるようになる。

5.4. 相関の強さ：★★★★

3 番目に相関の強いものについてみていこう。

まずは、全体を俯瞰する。

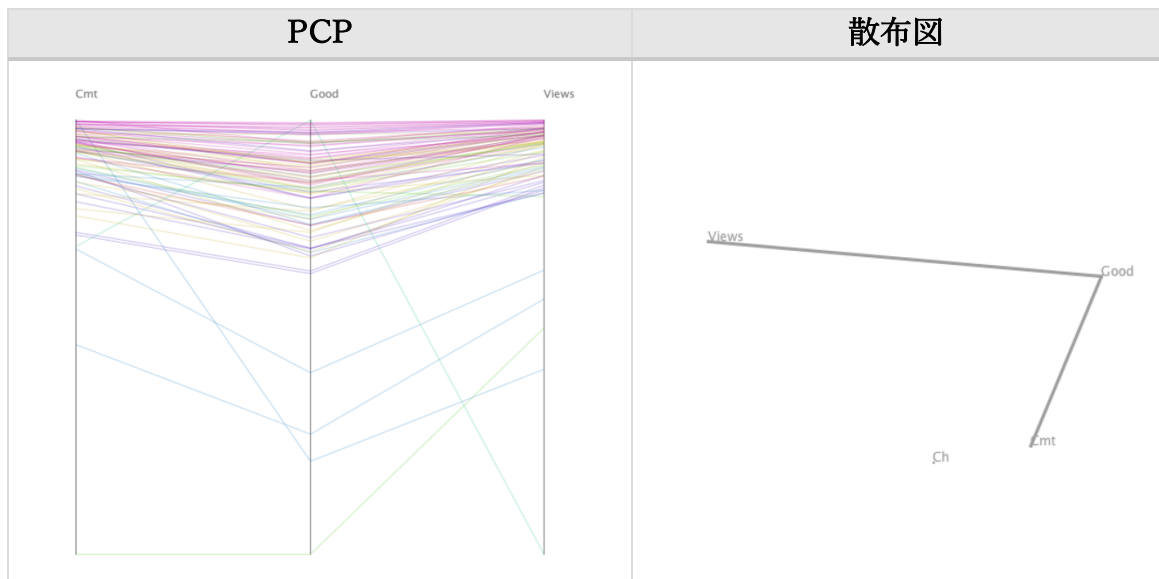




一見、当たり前の結果しか出ていないように思える。

しかし、ズームをしてよく目を通してみると、少し面白いことが発見できた。

オレンジ枠で囲まれた、[Cmt: Comment][Good][Views]について注目してみよう。



<PCP からわかること>

[Comment][Good][Views]にはやや強い正の相関がある。

<散布図からわかること>

[Views] ↔ [Good], [Good] ↔ [Comment] にはやや強い相関があるが、
[Views] ↔ [Comment] にはそれほど強い相関はない。

<考察>

多く見られているものは高評価の数が多く、高評価の数が多ければコメントも多くなるが、閲覧数が多いこととコメントの増加は直接的には結びつかない、
というのは、二つの理由が考えられるのではないだろうか。

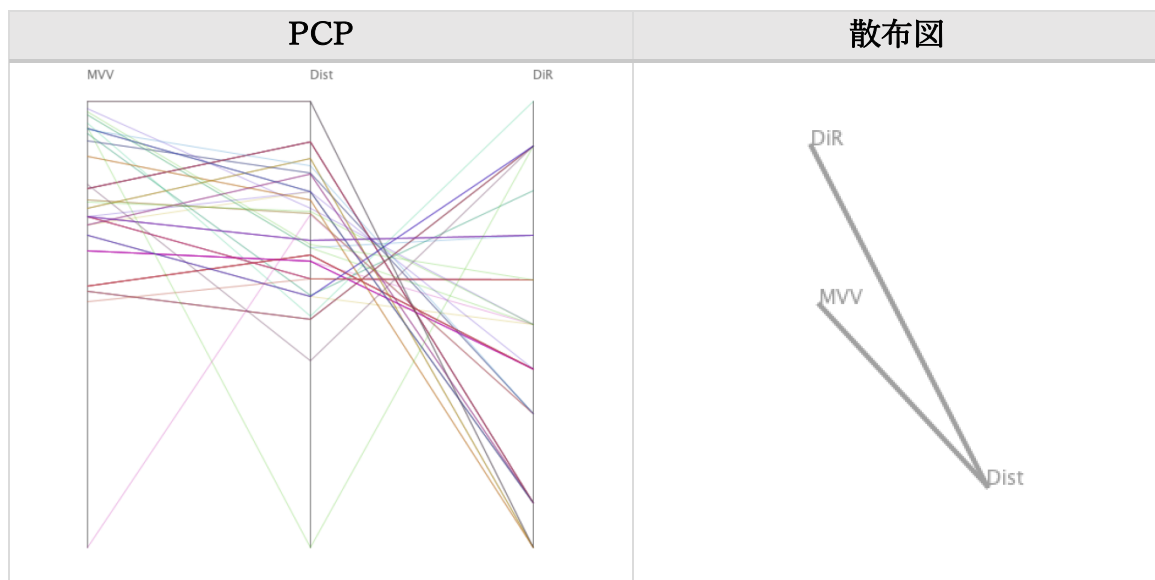
- ☆ ウォヌのチッケム動画には、コメントするほどの知識はない、すなわち大ファンではないが、サムネで惹かれて動画を閲覧し良かったために Good を押すのみに留める層が一定数いるため
- ☆ ウォヌのマイナーなチッケム動画まで見ている層は、コメントを活発に行うほどウォヌのファンであるから

これらについては、裏付ける根拠等は見つけられなかったが、考察として自然であろう。

ここからさらに意外な組み合わせの相関の強さを発見できるようになってくる。

5.5. 相関の強さ：★★

やや強い相関を持つ要素について見ていこう。



なんと、[MVV: MV Views][Dist: Distribution]([DiR: Distribution Rank])に強い相関があるらしい。

<PCP からわかること>

[MV Views]と[Distribution]にはやや強い負の相関がある。

<散布図からわかること>

[MV Views]と[Distribution], [Distribution]と[Distribution Rank]は同程度の強い相関がある。

<考察>

これはすなわち、“ウォヌの歌うパートが少ない方が、MV の再生回数が増える”ということである。非常に衝撃的な結果である。

SEVENTEEN に明るくない人にとっては、“ウォヌというメンバーは不人気なのだ”、と一蹴するかもしれないが、決してそんな事実はない。人気順が明らか

になる指標として、公式 Instagram([@saythename_17](#))における、アーティスト写真への”いいね”の数が挙げられるが、最新の「Sector; 17 NEW BEGINNING」においてはジョンハンに次ぐ二番目の”いいね”の多さであり、次の「Sector; 17 NEW HEIGHTS」においてもミンギュについて二番目に多い。したがって、人気メンバーである、“ウォヌの歌の配分が少ない曲の MV は再生回数が多い”という結果はファンにとっては驚愕なのだ。

このことについても、二つの憶測を立てよう。

- ✧ SEVENTEEN は明るい曲が人気(= 需要が高い)であり、声が低くラップを担当するウォヌはそういった曲では配分が少ないから、
- ✧ ウォヌが”Killing Part[*4]”を担当する楽曲の人気が高いから。

[*4]…Killing Part とは、曲の中で最も注目度の高い部分。通常 Killing Part を担当する場合、際立たせるために他のパートへの配分が少なくなる。

これらの仮説について、実際に検証してみる。

まず、以下は SEVENTEEN における MV 再生回数ランキング TOP5 の表である。

曲名	再生回数 (回)
울고 싶지 않아 (Don't Wanna Cry)	244,615,803
박수(CLAP)	113,561,370
Left & Right	107,749,091
HOT	105,743,161
Rock with you	86,738,672

(SEVENTEEN/[MUSIC VIDEO](#) より)

これらにおける、ウォヌの Line Distribution を見てみよう。

曲名	Dist (%)	Dist Rank (位)
울고 싶지 않아 (Don't Wanna Cry)	5.98	10
박수(CLAP)	7.84	7
Left & Right	9.00	4
HOT	7.13	9
Rock with you	5.95	10

やはり Left & Right をのぞいて、いずれにしてもウォヌの配分が少ないことがわかる。これはどうしてだろうか。まず、第一の仮説から考えてみる。

- ✧ 明るい曲が人気

“明るい曲”の定義について、[Spotify](#) を参考にすると、以下の通りである。

- テンポ: 楽曲の速さやリズムが速く、活発であること。

- キー: 高い音階や長調の音階が使用され、楽曲全体が明るく、陽気であること。
- メロディー: 軽快なリズムや単純で明快なメロディが特徴的であること。
- 歌詞: 明るく希望に満ちた歌詞が含まれること。
- 楽器: ギターやトランペットなど、明るい音色の楽器が使用されること

また、[Tunebat](#) 上のデータを参考にすると、人気曲の BPM/Energy は以下の通り。

曲名	BPM (bpm)	Energy (%)
울고 싶지 않아 (Don't Wanna Cry)	100	68
박수(CLAP)	100	92
Left & Right	150	83
HOT	112	78
Rock with you	164	82

ここでは BPM100 以上かつ Energy70%以上を”明るい”と定義すると、「울고 싶지 않아 (Don't Wanna Cry)」を除く人気曲のほとんどが明るい楽曲である。

したがって、

☆ 明るい曲が人気であるため、声が低くラッパー向きのウォヌは人気曲において歌うパートが少ない

という仮説は正しそうである。

第二の仮説として、

☆ ウォヌが Killing Part を担当している

これについて、先ほど”明るい曲”に当てはまらなかった「울고 싶지 않아 (Don't Wanna Cry)」を見てみようと思う。

[Music Video](#) を見て貰えばわかると思うが、最初の歌の入りがウォヌである。もうすでに Killing Part を担当しているといっても過言ではない。さらに、2:33~からのウォヌのパート後、全ての音が止む。これは完全なる Killing Part である。実際にライブでは、この後をファンが歌うこともあるくらい、この曲において象徴的なパートである。

したがって、第二の仮説、

☆ 人気曲においてウォヌが Killing Part を担当しているため、歌のパートが少ない

も正しいといえよう。

まとめとして、

“ウォヌの歌のパートが少ない歌ほど、再生回数が多い”という相関が存在し、

その理由としては

✧ ウォヌの声質には合わない、明るい曲が人気だから。

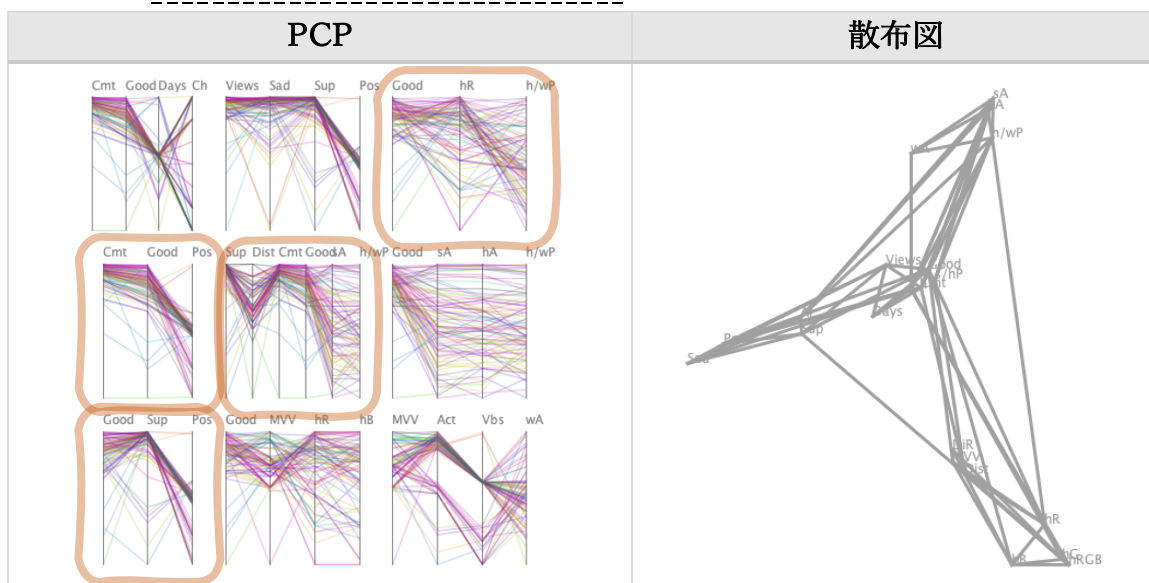
✧ ウォヌが人気曲の Killing Part を担当しているため、必然的に配分が少なくなるから。

が挙げられる。

5.6. 相関の強さ：★

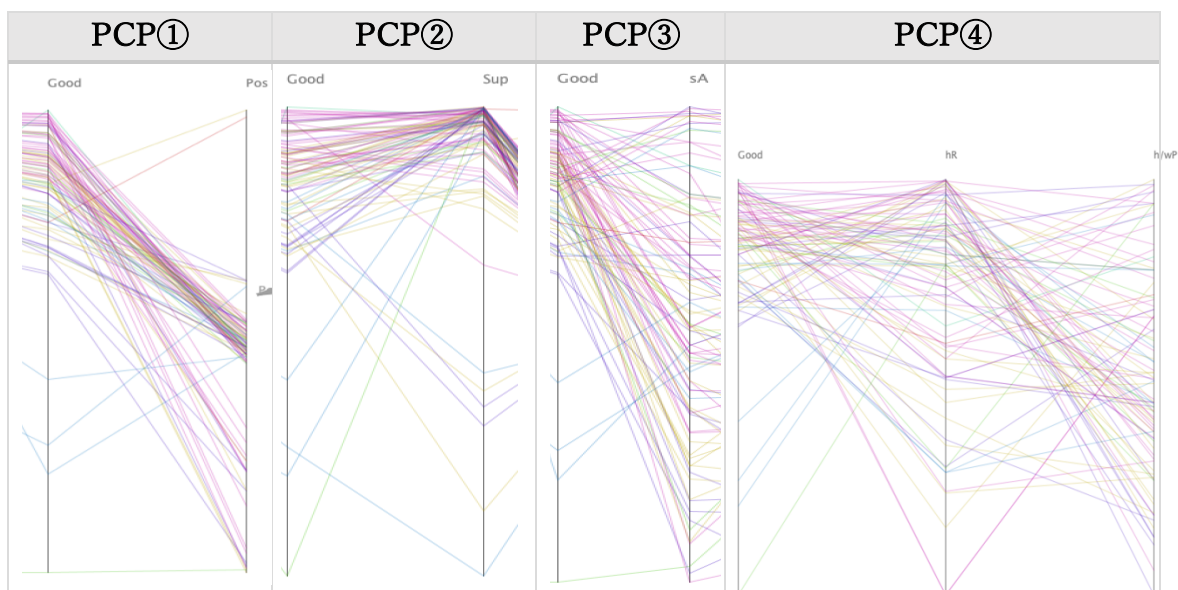
最後に、相関がある程度強い要素について見てみよう。

最も面白い結果を得ることができた。



オレンジ枠で囲まれた部分について、詳しく見てみる。

ここからは、散布図については複雑に絡み合ってしまう、観察しにくいので PCP について考察していくことにする。



<PCP①からわかること>

[Good][Pos: Positive]にはある程度強い負の相関がある。

<PCP②からわかること>

[Good][Sup: Surprised]にはある程度強い負の相関がある。

<PCP③からわかること>

[Good][sA: skinArea]にはある程度強い正の相関がある。

しかし、skinArea が大きい部分には、負の相関もみられる。

<PCP④からわかること>

[Good][hR: hairRed]にはある程度強い負の相関がある。

[hR: hairRed][h/wP: head/wonwoo Percent]にはある程度強い負の相関がある。

→[Good][h/wP: head/wonwoo Percent]には正の相関がある。

<考察>

ウォヌのチッケムにおいて、より多くの高評価を得るには、

- ✧ サムネにはポジティブではない表情のものを選ぶ
- ✧ サムネには驚いていない表情のものを選ぶ
- ✧ サムネは程よく肌を見せつつ、見せすぎないようにする。
→ 顔ははっきりと見えるが、短髪すぎないものを選ぶ？
- ✧ サムネには髪の色を抑えたものを選ぶ
- ✧ サムネは、全身を写すよりも顔が見えやすいものを選ぶ

とよい、ということが推察される。

それぞれの妥当性について、ウォヌ本人のインスタグラム([@everyone_woo](#))の”いいね数”をもとに検討してみる。

以下が、カテゴリと”いいね”数の平均値の表である。

カテゴリ	<u>Non-Expressive</u>	<u>Expressive</u>	<u>Non-Bangs</u>	<u>Bangs</u>
いいね(万件)	128.748	114.0083	136.12	120.932
カテゴリ	<u>Non-Red</u>	<u>Red</u>	<u>Whole body</u>	<u>Upper body</u>
いいね(万件)	131.1	109.725	123.152	125.6667

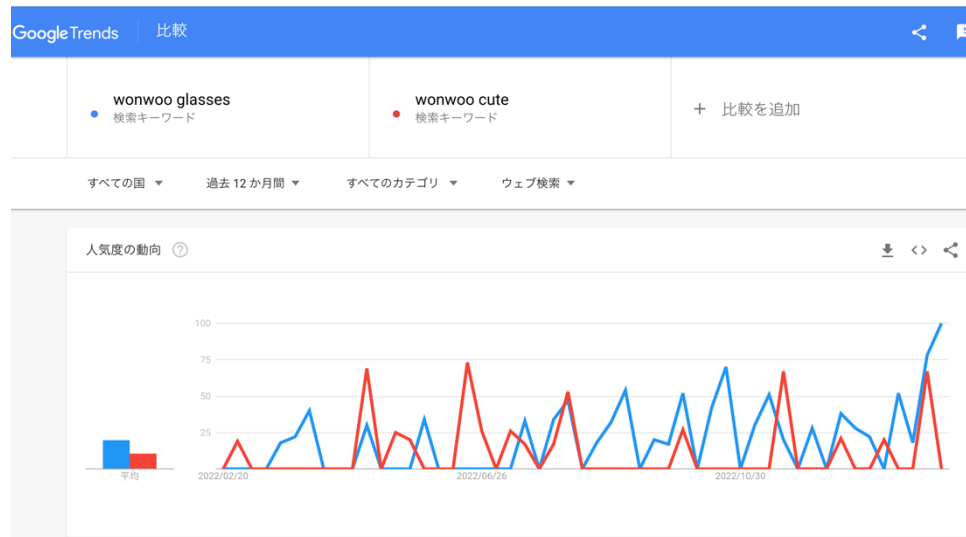
(表 5.6)

- ✧ ウォヌはポジティブでない表情の方が好まれる
- ✧ ウォヌは驚いていない表情の方が好まれる

これらについては、”表情が豊かか否か”でカテゴリ分けした。表 5.6 を見ると、チッケムのサムネの相関と同じように、表情が豊かではない方が”いいね”数が多い、という結果になった。

すなわち、ウォヌは無表情の方が需要があるということである。これはアイドルとしては意外な結果ではあるが、二重で顔が丸く、可愛い顔立ちというよりは、一重で顎がしっかりとした端正な顔立ちであるので、無表情の方がその美しさを堪能できる、という点で頷ける。

また、調べていくうちに以下のような結果も出た。



[ここからアクセス可能である。](#)

“wonwoo glasses”の方が、“wonwoo cute”よりも需要が高いことがわかった。このことは、ファンがウォヌに求めるものとは、可愛い姿よりもカッコよく聡明な姿であることを表しているだろう。

☆ 程よく肌を見せつつ、見せすぎないようにする
→ 短髪は印象が良くないのか？

これは、表 5.6 を見ると、前髪のないウォヌの写真のほうが“いいね”が多くきているので、間違った考察であったと言える。確かに、ウォヌのチッケムの中で最も再生されているものと二番目に多く再生されているものは、共に前髪のないものであった。したがって、前髪はない方が人気であり、[hA: headArea]よりも[sA: skinArea]の方が相関が強いことを鑑みると、衣装の違いで高評価に差がついている、という可能性がより高そうである。

☆ 髪の色が赤みの少ない方が良い

これについては、表 5.6 から明らかなように、ウォヌは黒髪需要の方が高そうである。確かに、チッケムのコメント欄でも以下のようなものも多く見られる。また、そういったコメントに対する“いいね”の数も非常に多い。



3 年前 (編集済み)

He is seriously handsome. A lot of people are upset that he doesn't change his hair color often but like, ladies, he looks superior with his black hair T.T

👍 6341 🗨️ 返信

▼ 39 件の返信

@MCOUNTDOWN_2019.8.8/「세븐틴 원우 직캠 4K 'HIT」のコメント欄より

☆ 全身よりも顔を写す方が良い

表 5.6 を参照すると、このことについても正しそうである。彼はモデルではなくアイドルなので当然な結果ではある。

6. 感想

今回、公開されているデータ等を扱わず、自分で一から csv データを作成したので非常に大変だった。特に、顔や頭部、身体の面積を測定する作業は 10 時間以上かかり、心が折れかけた。また、YouTube API を初めて用いて Python を動かし、動画情報の取得を試みたことも、かなり苦労した。しかし、推しへの[あいのちから](#)で何とか乗り切り、どのようなサムネイルのチェックが伸びやすいのか、といった素朴な疑問から、ファンが求めるウォヌ像の輪郭まで捉えることができ、本当に嬉しかった。

今回調べてみて驚いたこととしては、ウォヌのパートが少ないものは MV の再生回数が多という相関があることと、黒髪の人気が高く、前髪はない方が良いということである。私としては、普段黒髪だからこそ、時々見られるハイトーンのウォヌが好きだったので驚いた。

唯一後悔しているのは、Twitter や YouTube のコメント欄からファンの方々のコメントやツイートを取得し、“WONWOO”や”ウォヌ”と共にどのような言葉が使われているのかを調べる形態素解析ができなかったことだ。今後、またデータ分析をする機会があれば、それまでに Python や R, Mecab をマスターして活用できるようになりたい。

以下、作成した csv データの一部である。先生のおっしゃる通り、好きなものについての分析にしたため、知りたいことが山ほどあって候補を絞るのが大変だった。

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
1	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric
2	Views	Cnt	Good	Days	MVV	Nr	Hp	Sad	Ar	Sup	Pos	Act	Vbs	hR	hG	hB	hRGB	Dist	DR	sA	hA	s/P	wA	h/wP
3	396921	481	23977	6	67859561	28.24	7.39	6.55	5.22	52.6	51.52	75.54	77.07	149	84	80	313	7.83	7	18739	35347	53.0144001	88678	39.4154642
4	106273	180	8169	10	86558379	80.84	2.75	4.63	1.15	10.64	61.42	36.82	48.24	11	8	35	54	7.31	8	26272	62734	41.8784072	79436	78.9742686
5	321546	294	19489	7	78283242	69.06	7.11	1.07	9.25	13.52	59.76	38.74	48.5	45	25	26	96	8.35	8	20100	29570	67.9742983	71559	41.322545
6	315674	213	16941	6	53108161	78.58	11.69	1.63	0.75	7.35	61.1	37.19	48.29	135	116	159	410	3.88	12	37518	62700	59.8373206	78263	80.114858
7	64314	64	4704	-245	20839007	52.75	42.02	0.88	1.39	2.95	57.45	41.4	48.85	116	106	130	352	5.3	12	18000	32219	55.8676557	66489	48.4576396
8	176116	96	10953	13	53108161	73.88	11.76	0.88	11.87	1.62	60.44	37.95	48.39	138	132	132	402	3.88	12	21077	44892	46.9504589	63901	70.2524217
9	36060	41	2829	21	53108161	81.44	7.07	3.4	7.74	0.36	61.51	36.72	48.23	162	139	155	456	3.88	12	31194	51300	60.8070175	76076	67.4325614
10	178446	312	13571	338	72721551	0.52	98.65	0.81	0.01	0.01	92.25	59.47	100	181	184	197	522	5.3	12	10793	28064	38.4557628	95704	29.325838
11	523817	1099	33415	7	107669450	1.4	96.45	1.94	0.02	0.19	92.28	59.26	100	0	0	0	0	9	4	39286	68066	57.7175095	70808	96.1275562
12	250210	306	15017	135	24063196	95.43	1.74	1.45	0.05	1.34	63.48	34.44	47.93	32	2	80	114	8.3	5	40203	72816	55.2117666	115278	63.165565
13	68429	106	5743	198	107669450	67.32	30.89	1.58	0.18	0.02	59.51	39.02	48.54	7	6	37	50	9	4	41490	70845	58.5644717	105492	67.1567512
14	963471	587	28840	127	24063196	71.22	4.61	2.07	10.75	11.36	60.06	38.39	48.45	68	75	117	260	8.3	5	16286	32812	49.6342801	95992	55.061082
15	70584	110	6027	190	107669450	81.97	4.64	7.29	3.57	2.54	61.58	36.64	48.22	48	7	65	120	9	4	23696	54142	41.7653921	99611	54.3534349
16	128322	267	11419	200	27541649	0.27	97.77	0.81	0.03	1.13	92.86	59.39	100	56	48	97	201	4.78	10	27104	55725	48.6388515	88786	62.7632735
17	349822	456	20962	16	27541649	68.91	24.38	2.01	0.41	0.3	59.74	38.77	48.5	20	36	96	152	4.78	10	24054	41133	58.4785938	72000	57.1291667
18	189854	210	13727	11	105314852	85.1	0.75	13.18	0.27	0.71	62.03	36.12	48.15	24	25	45	94	7.13	9	15781	24572	64.2235064	61536	39.9310972
19	428417	781	28083	3	86558379	77.79	20.11	1.32	0.49	0.3	60.99	37.32	48.31	2	0	11	13	7.31	9	38492	80363	47.8976644	96100	83.6243496
20	334489	375	23970	3	105314852	25.52	3.35	4.04	0.12	66.56	51.93	82.32	84.25	0	10	22	32	7.13	9	33505	48063	69.7105882	74182	64.8081228
21	277888	254	17005	16	59070816	85.55	2.83	10.54	0.65	0.43	62.09	36.05	48.14	0	0	34	34	5.95	10	29448	55106	53.4388827	77130	71.4456113
22	506838	1079	32946	7	63831815	26.73	72.74	0.2	0.1	0.23	81.89	56.98	88.88	22	25	70	117	4.34	13	29052	54013	53.7870513	68363	79.0091131
23	363438	655	28605	2	105314852	76.64	5.16	1.21	2.4	14.59	60.83	37.5	48.33	13	6	14	33	7.13	9	9109	21988	41.4271421	83271	26.4053512
24	1068607	4164	64264	13	107669450	1.19	88.74	0.1	0.01	0.01	94.94	60.48	100	44	29	26	01	0	4	45684	64964	71.6216681	89764	71.6216681