

2016年 S-PLUS & Visual R Platform 学生研究奨励賞

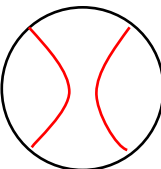
プロ野球における日本人選手の役割 ～2016年度と2015年度のチーム状況の違い～

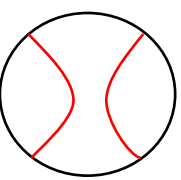
目次

1. 研究目的
 2. 使用データ概要
 3. 研究の流れ
 4. 分析結果
 5. まとめ、今後の課題
- 参考文献-

東海大学 情報通信学部
経営システム工学科

松家 大士





研究目的

セ・リーグ

ヤクルトスワローズ
読売ジャイアンツ
阪神タイガース
広島カープ
中日ドラゴンズ
DeNAベイスターズ

NPB
(日本野球機構)

パ・リーグ

ソフトバンクホークス
日本ハムファイターズ
千葉ロッテマリーンズ
西武ライオンズ
オリックスバファローズ
楽天イーグルス

表1 2015年度セ・リーグ順位

セ・リーグ	チーム
1位	ヤクルト
2位	巨人
3位	阪神
4位	広島
5位	中日
6位	DeNA

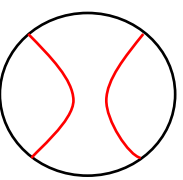
表2 2016年度セ・リーグ順位

セ・リーグ	チーム
1位	広島
2位	巨人
3位	DeNA
4位	阪神
5位	ヤクルト
6位	中日

2015年と2016年何が違うの??

目的

2015年度と比較して
2016年度は何が影響しているのか?



使用データ概要

・提供元

NPB公式記録^[1]

・データ概要

日本プロ野球で行われた各対戦カードの試合結果 計143試合

・期間

2015年度、2016年度レギュラーシーズン(交流戦を含む)

・使用範囲

全12球団

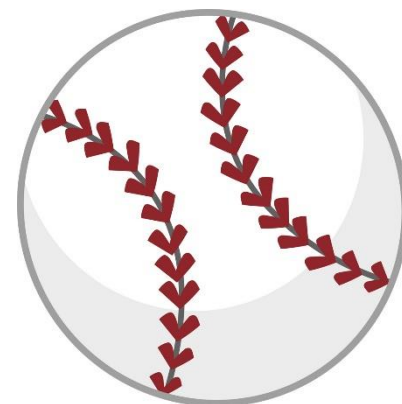
・分析対象

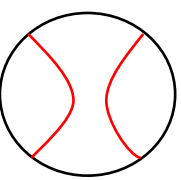
2015年度: 658人の打撃データ

2016年度: 680人の打撃データ

・データ項目

安打, 二塁打, 三塁打, 本塁打, 打点, 盗塁, 犠打, 四球, 死球, 三振, 併殺打, 得点, 塁打, 犠飛, 故意四球





研究の流れ

NPBのデータを参考に
両年度の打者データの収集

主成分分析の結果をグラフにプロットする。縦軸をアベ
レージヒッター、横軸をパワーヒッターと定義

データ
収集

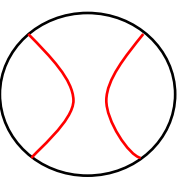
主成分分析

プロット

チームの
現状把握

11個の変数を使う。
(安打,二塁打,三塁打,本塁打,打点,
盗塁,犠打,四球,死球,三振,併殺打)

主成分プロットしたデータの外側を
円で描き、その円の大きさを違うことを確認
共分散構造分析により、背景を探る



分析結果(2015年度)

表3 プロ野球12球団2015年度打者成績
主成分分析による分布表

	成分	
	1	2
安打	0.979	0.032
二塁打	0.953	-0.090
三塁打	0.589	0.501
本塁打	0.847	-0.372
打点	0.954	-0.218
盗塁	0.643	0.546
犠打	0.412	0.715
四球	0.943	-0.057
死球	0.741	-0.125
三振	0.944	-0.045
併殺打	0.862	-0.201

主成分分析で使用する指標

安打、二塁打、三塁打、本塁打、打点、盗塁、犠打、四球、死球、三振、併殺打

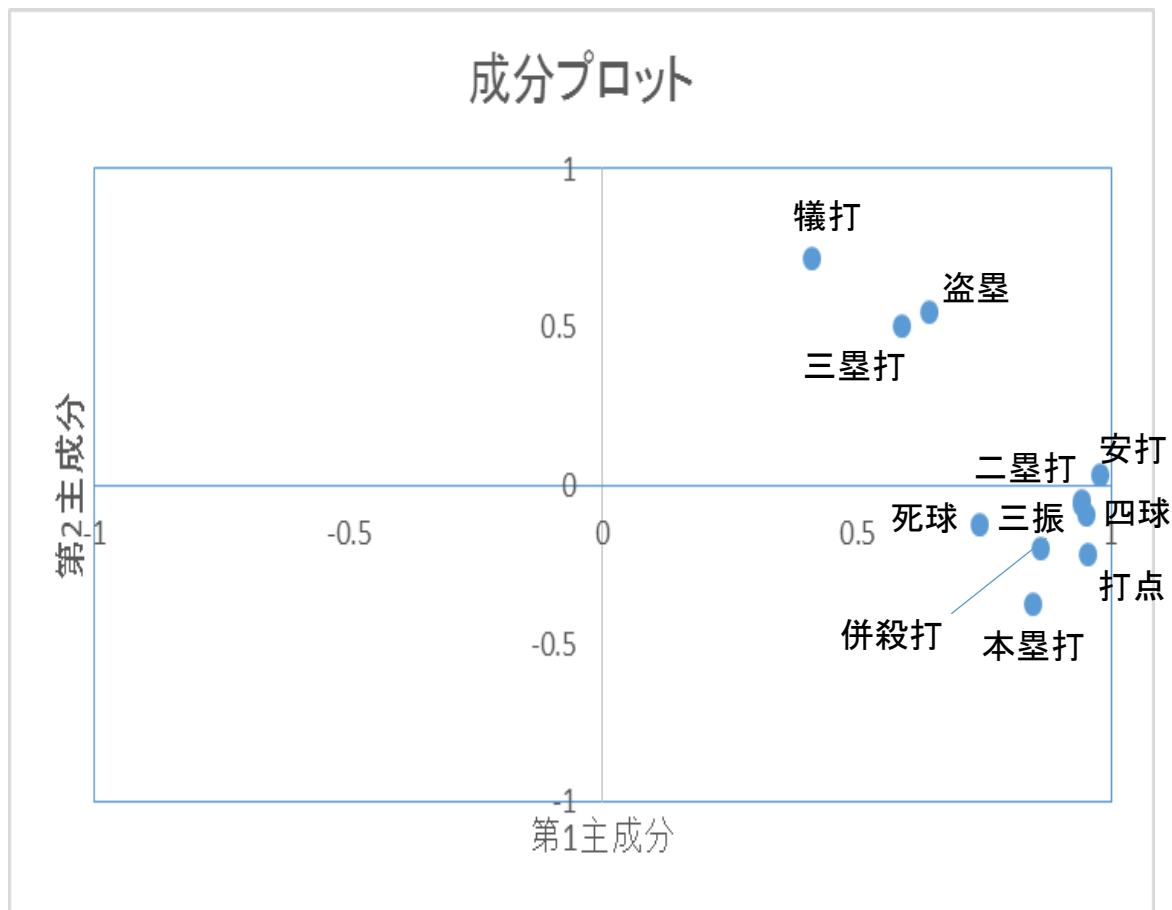
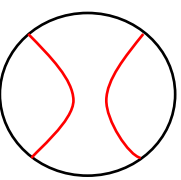


図1 プロ野球12球団2015年度打者成績の主成分分析



分析結果(2016年度)

表4 プロ野球12球団2016年度打者成績主成分分析による分布表

	成分	
	1	2
安打	0.979	0.022
二塁打	0.959	-0.082
三塁打	0.688	0.411
本塁打	0.819	-0.436
打点	0.952	-0.241
盗塁	0.619	0.514
犠打	0.384	0.688
四球	0.938	0.004
死球	0.771	0.022
三振	0.944	-0.019
併殺打	0.84	-0.269

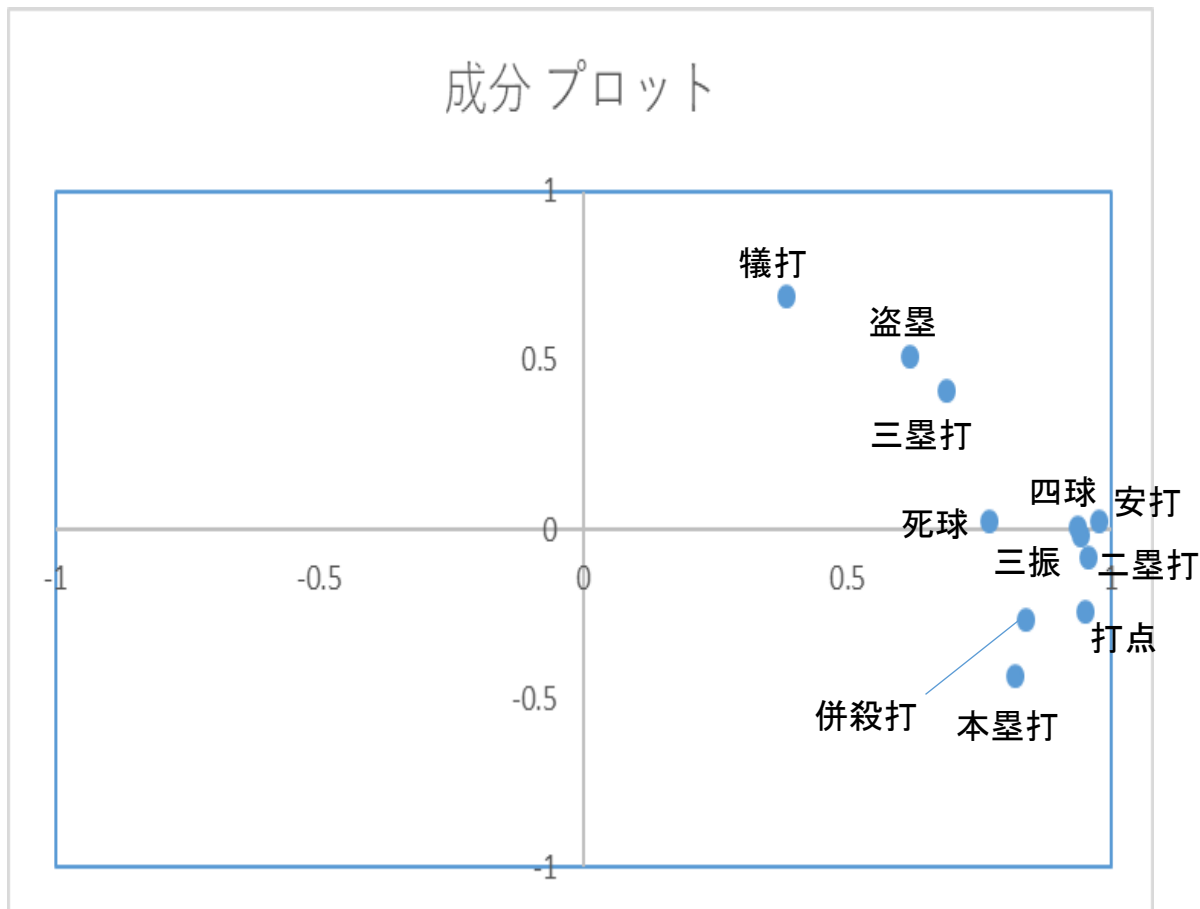
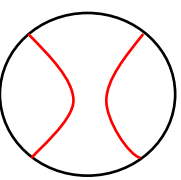


図2 プロ野球12球団2016年度打者成績の主成分分析

主成分分析で使用する指標

安打、二塁打、三塁打、本塁打、打点、盗塁、犠打、四球、死球、三振、併殺打



分析結果(主成分分析)

表5 プロ野球12球団2015年度打者成績の主成分分析による成分表

成分	初期の固有値		
	合計	分散(%)	累積(%)
1	7.5	68.130	68.130
2	1.3	11.977	80.107
3	0.6	5.760	85.867
4	0.5	4.234	90.101
5	0.4	3.959	94.059
6	0.2	2.240	96.300
7	0.2	1.458	97.757
8	0.1	.890	98.648
9	0.1	.876	99.523
10	0.0	.306	99.829
11	0.0	.171	100.000

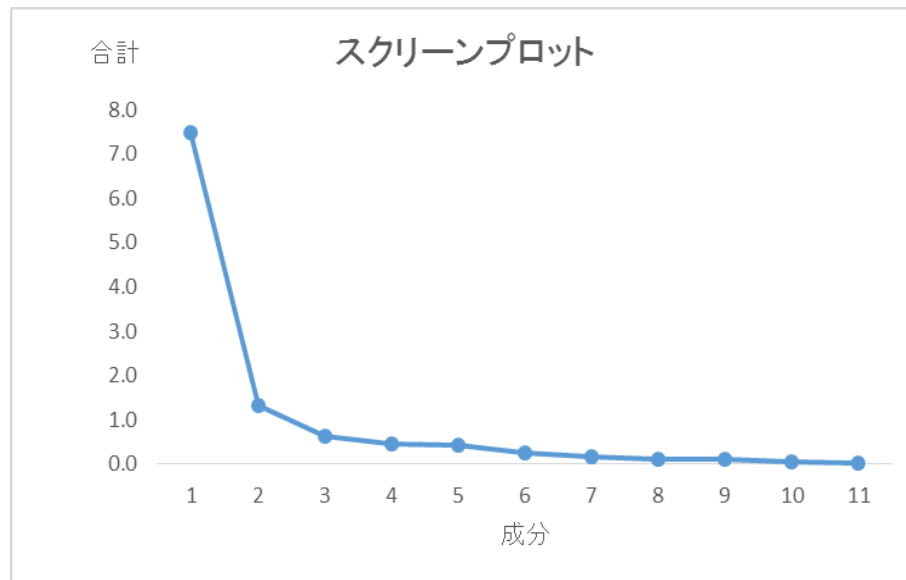
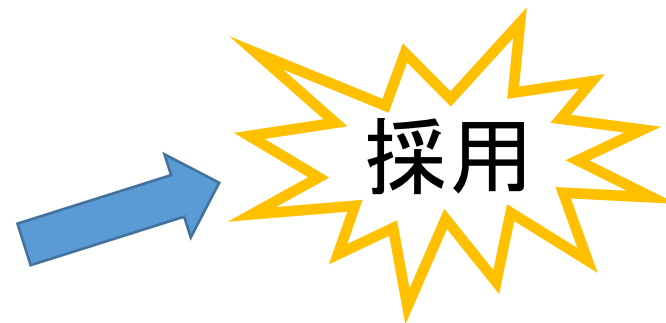
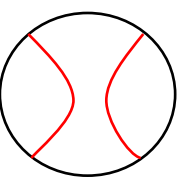


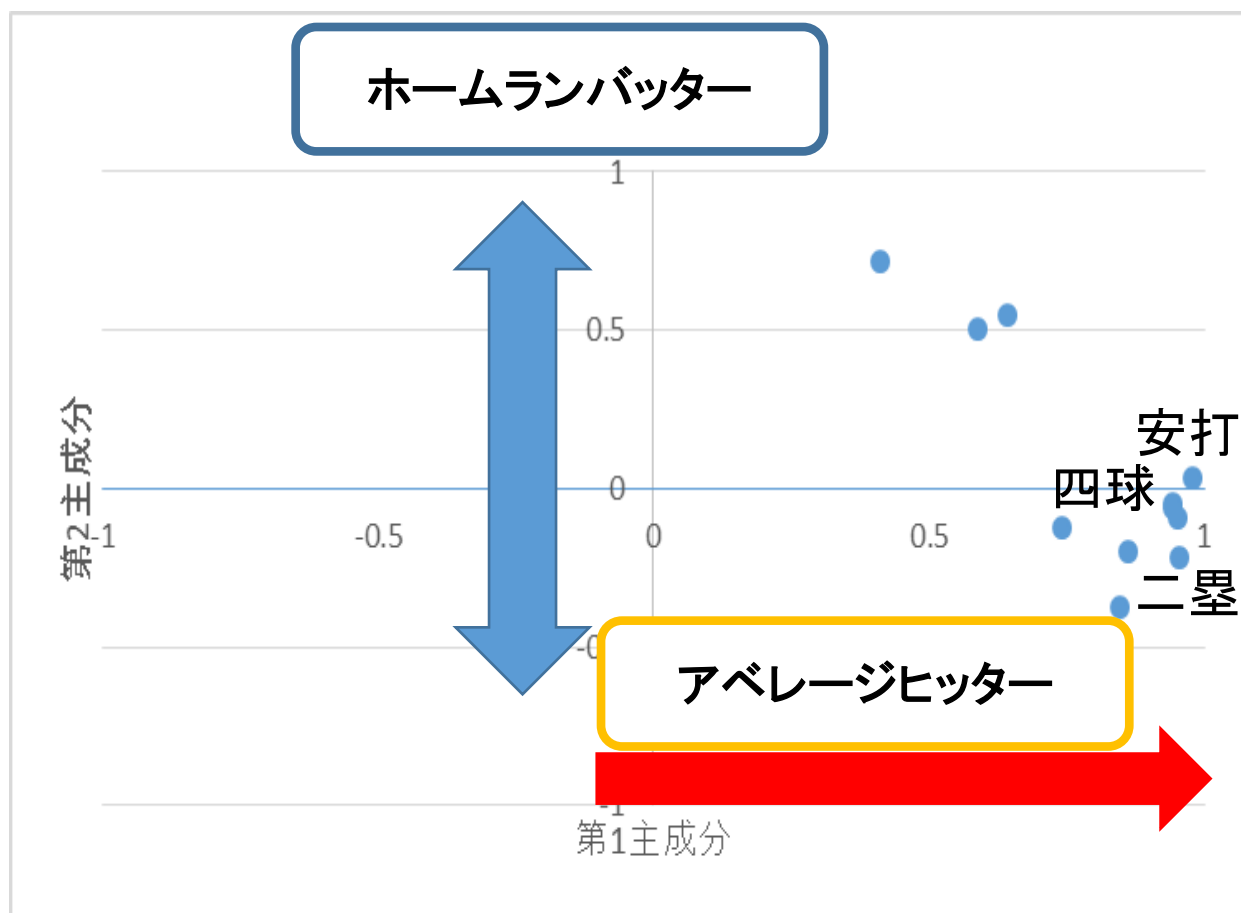
図3 2015年度打者成績の主成分分析による成分図

累積寄与率が第2主成分で80%以上ある
第3主成分では分散約5%





分析結果



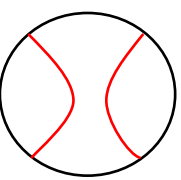
・第一主成分
右に行くほどヒット性の
当たりを打つ人が多い

アベレージヒッター

・第二主成分
上に行くほど
盗塁する人が多い
下に行くほど本塁打が多い

ホームランバッター

図4 プロ野球12球団2015年度打者成績の主成分分析による成分分布図

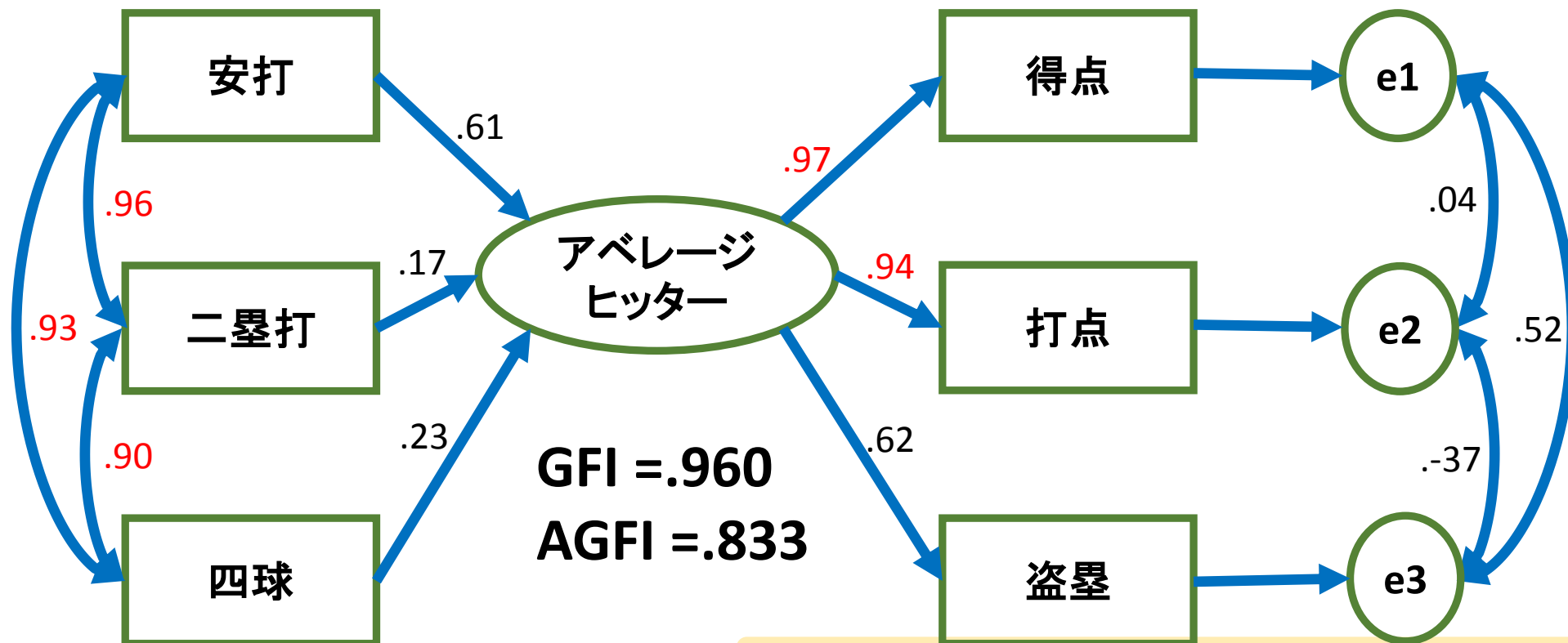


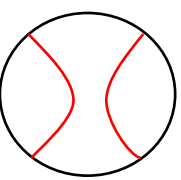
分析結果（共分散構造分析）

- ✓ アベレージヒッターが得点に大きく関連している
- ✓ アベレージヒッターが打点に絡むことが多い
- ✓ 安打、二塁打、四球の共分散が高い

* アベレージヒッターとは主成分分析で出た第1主成分の結果を入れた

* eは誤差変数を表す



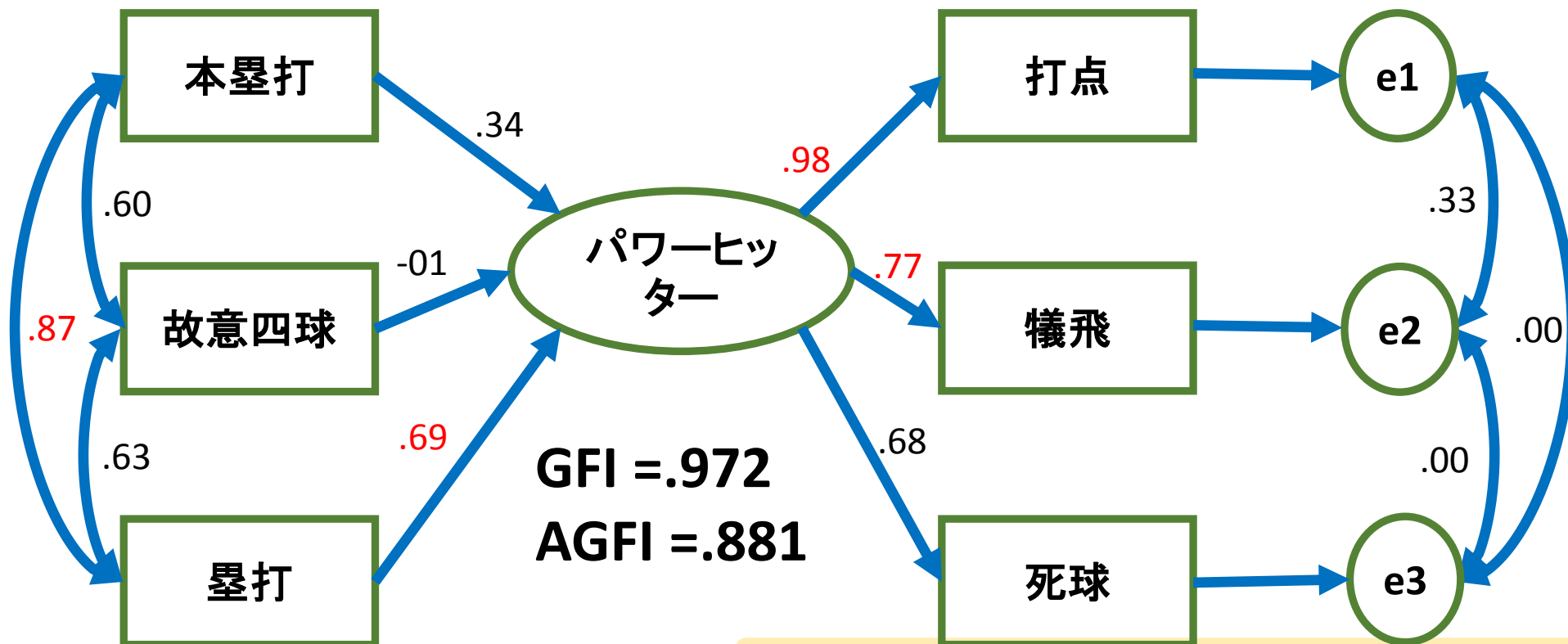


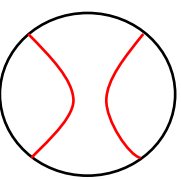
分析結果（共分散構造分析）

- ✓ パワーヒッターが打点に大きく関連している
- ✓ パワーヒッターが犠飛に絡むことが多い
- ✓ 本塁打と塁打の共分散が高い

* パワーヒッターとは主成分分析で出た第2主成分の結果を入れた

* eは誤差変数を表す





主成分分析結果(広島)

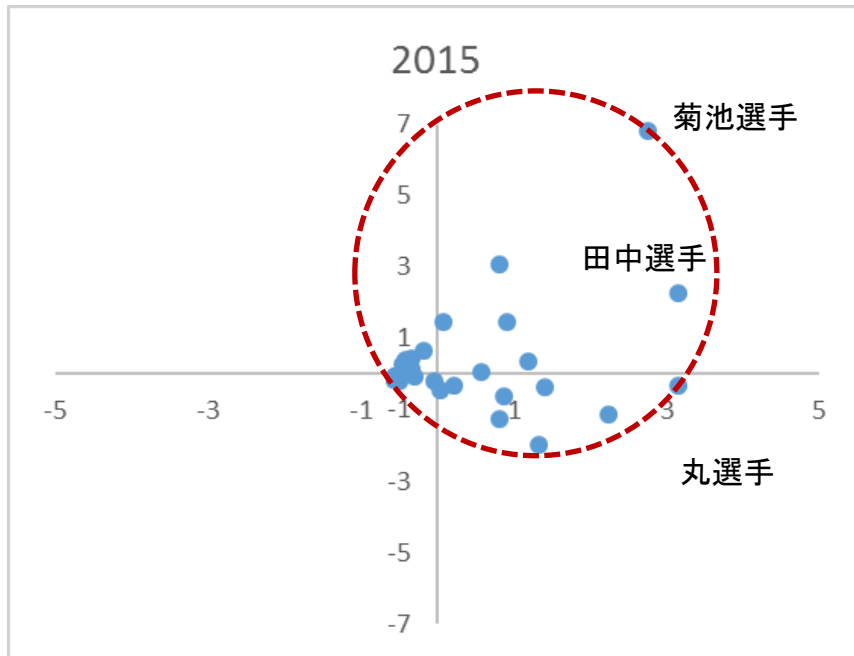


図6 広島2015年度打者成績の主成分分析による分布図

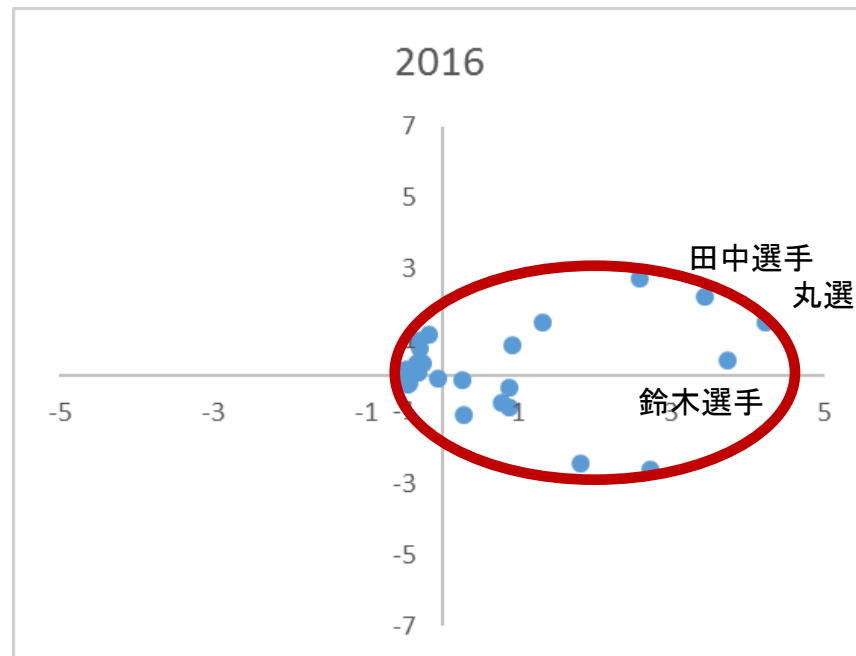
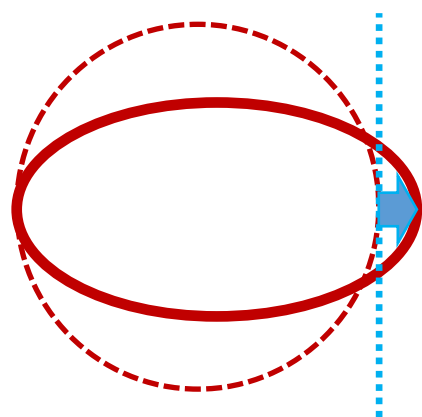


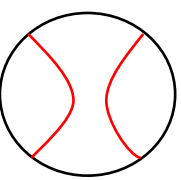
図5 広島2016年度打者成績の主成分分析による分布図



- 円の大きさは2015年度の方が大きい
- **プロットデータが横に分散している**



プロットデータが横に伸びることに勝率のカギがある！？！？



主成分分析結果(広島)

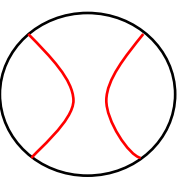
2015	横軸	縦軸
選手	FAC1_1	FAC2_1
丸 佳浩	3.15635	-0.3195
田中 広輔	3.15171	2.2609
菊池 涼介	2.75932	6.79362
新井 貴浩	2.23764	-1.1468
會澤 翼	1.40015	-0.3454
エルドレッド	1.32806	-1.9872
梵 英心	1.19963	0.36256
鈴木 誠也	0.92029	1.44936
シアーホルツ	0.87364	-0.6052
平均	1.7338	

表6 2015年度広島主成分得点表

2016	横軸	縦軸
選手	FAC1_1	FAC2_1
丸 佳浩	4.22708	1.49121
鈴木 誠也	3.7367	0.45358
田中 広輔	3.43072	2.2157
新井 貴浩	2.72846	-2.6089
菊池 涼介	2.58027	2.71972
エルドレッド	1.80529	-2.4676
安部 友裕	1.31059	1.48379
石原 慶幸	0.92014	0.8701
松山 竜平	0.88901	-0.8572
平均	2.17515	

表7 2016年度広島主成分得点表

- 2015年度と2016年度で平均が0.5違う！！
- 横軸が伸びれば勝率も上がる！？



主成分分析結果 (DeNA)

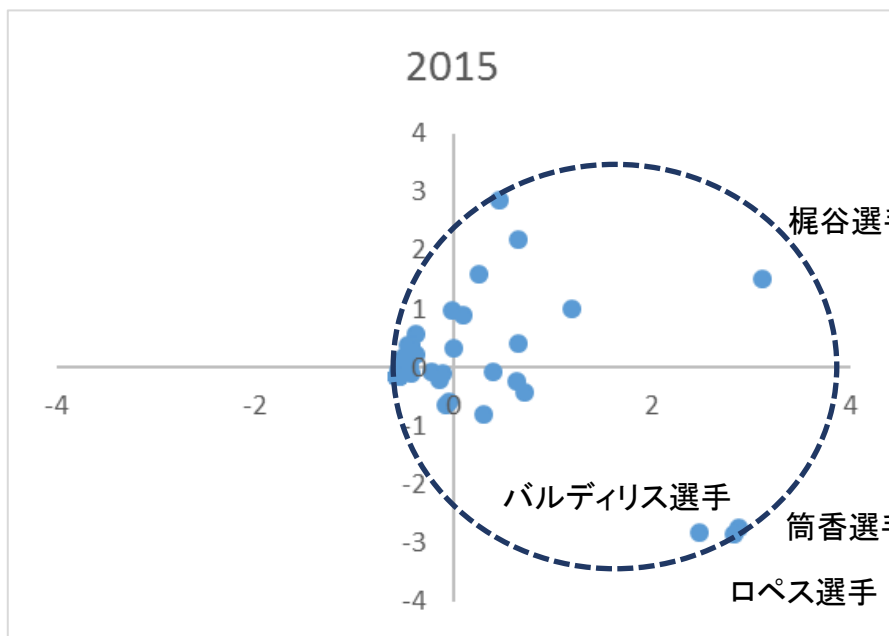


図8 DeNA2015年度打者成績の主成分分析による分布図

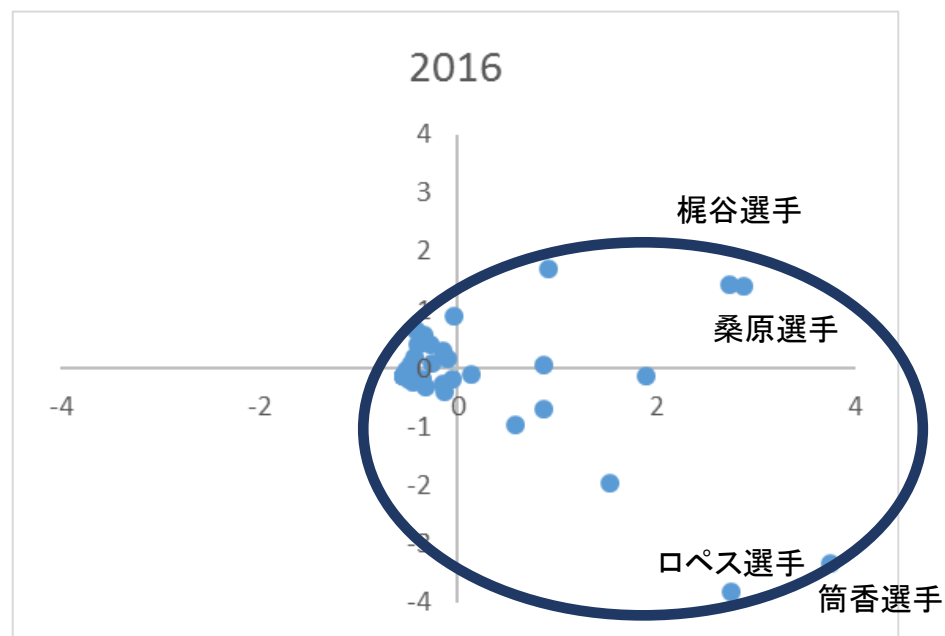
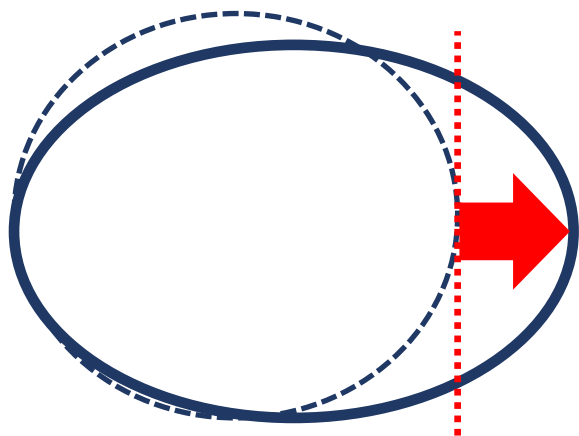


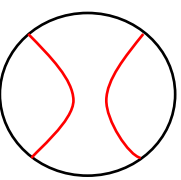
図7 DeNA2016年度打者成績の主成分分析による分布図



- 2016年度の方が円が大きくなっている
- **プロットデータが横に分散している**



プロットしたデータが横に伸びる ことに勝率のカギがある！！



主成分分析結果 (DeNA)

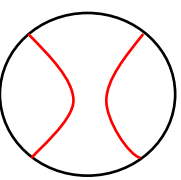
2015	横軸	縦軸
選手	FAC1_1	FAC2_1
梶谷 隆幸	3.10932	1.52008
筒香 嘉智	2.88201	-2.7344
ロペス	2.83427	-2.8503
バルディリス	2.4835	-2.807
石川 雄洋	1.19614	0.99247
嶺井 博希	0.71719	-0.4352
白崎 浩之	0.66259	2.18485
荒波 翔	0.6568	0.40122
倉本 寿彦	0.6354	-0.229
平均	1.68636	

表8 2015年度DeNA主成分得点表

2016	横軸	縦軸
選手	FAC1_1	FAC2_1
筒香 嘉智	3.754	-3.3379
桑原 将志	2.88624	1.41575
ロペス	2.75712	-3.823
梶谷 隆幸	2.7337	1.43061
倉本 寿彦	1.88873	-0.1379
宮崎 敏郎	1.53483	-1.9478
石川 雄洋	0.9161	1.71049
エリアン	0.87128	0.0657
戸柱 恭孝	0.86409	-0.6874
平均	2.0229	

表9 2016年度DeNA主成分得点表

- 去年と今年で平均が0.4違う！！
- 広島と同じように横軸が伸びている！



主成分分析結果(ヤクルト)

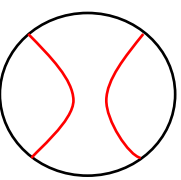
2015	横軸	縦軸
選手	FAC1_1	FAC2_1
山田 哲人	4.79182	-0.4121
畠山 和洋	2.92868	-3.5031
川端 慎吾	2.69487	-1.3398
雄平	2.47127	0.26846
中村 悠平	1.65132	0.42319
大引 啓次	1.29591	0.28358
デニング	0.84524	-1.2411
上田 剛史	0.73591	1.96765
比屋根 渉	0.61932	0.89493
平均	2.00381	

表10 2015年度ヤクルト主成分得点表

2016	横軸	縦軸
選手	FAC1_1	FAC2_1
山田 哲人	4.55224	-1.2021
バレンティン	3.17882	-4.3236
坂口 智隆	2.31981	1.85618
雄平	1.46453	-1.0983
川端 慎吾	1.45195	-0.7577
中村 悠平	1.36431	0.44017
大引 啓次	1.31265	-0.0338
西浦 直亨	0.85159	-0.1507
西田 明央	0.83961	-0.495
平均	1.92617	

表11 2016年度ヤクルト主成分得点表

- FAC1-1の平均が2015年度が2016年度より下がっている。
- 山田、雄平、川端、畠山選手はケガで離脱している時期がある。



まとめ

まとめ

- 横軸に円が伸びている方が勝率が高い
- 来年度、補強するなら円が横に伸びる方がよい
- ケガなどで出場回数が減るとチームの勝率が下がる

	2015	2016
チーム名	FAC1_1	FAC1_1
広島	1.73380	2.17515
DeNA	1.68636	2.02290
巨人	1.70585	1.84998
ヤクルト	2.00381	1.92617
阪神	1.71573	1.76626
中日	1.63888	1.77127

表14 各チームの主成分得点

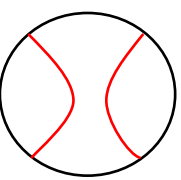
打点順位		
1	筒香 嘉智	D
2	山田 哲人	ヤ
3	新井 貴浩	広
4	バレンティン	ヤ
5	鈴木 誠也	広
5	ロペス	D
7	丸 佳浩	広

表12 2016年度打点ランキング

本塁打順位		
1	筒香 嘉智	D
2	山田 哲人	ヤ
3	ロペス	D
4	バレンティン	ヤ
5	鈴木 誠也	広

表13 2016年度本塁打ランキング

広	広島
D	DeNA
ヤ	ヤクルト



今後の課題

今後の課題

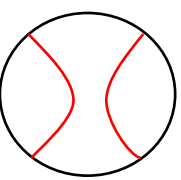
- 3年以上のデータ使用して精度を上げる
- ピッチャーのデータを使用する
- 費用対効果の検証

最多勝			防御率		
1	野村 祐輔	広	1	菅野 智之	巨
2	ジョンソン	広	2	ジョンソン	広
3	メッセンジャー	神	3	野村 祐輔	広
4	山口 俊	D	4	田口 麗斗	巨
5	田口 麗斗	巨	5	岩貞 祐太	神
5	岩貞 祐太	神	6	メッセンジャー	神
5	黒田 博樹	広	7	黒田 博樹	広
8	菅野 智之	巨	8	石田 健大	D
8	石田 健大	D	9	藤浪 晋太郎	神
8	内海 哲也	巨	10	井納 翔一	D

表15 セ・リーグ勝利数、防御率ランキング

順位	チーム	年棒合計
1位	巨人	32億円
2位	阪神	24億円
3位	ヤクルト	21億円
4位	広島	18億円
5位	中日	16億円
6位	DeNA	14億円

表16 セ・リーグチーム別年棒順位^[5]



参考文献

[1] 日本野球機構 (<http://www.npb.or.jp/>)

最終閲覧日; 2016年10月25日

[2] コトバンク (<https://kotobank.jp/>)

最終閲覧日; 2016年10月25日

[3] 木下栄蔵

「阪神はなぜ優勝したか(スポーツデータII)」

オペレーションズ・リサーチ, 51(2), 111-115, 2006

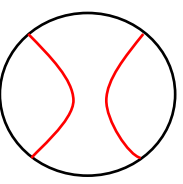
[4] 日刊スポーツ: 記事2016年4月26日

(<http://www.nikkansports.com/baseball/news/1637560.html>)

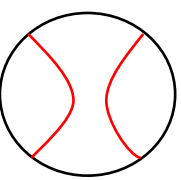
最終閲覧日; 2016年10月25日

[5] 鍋谷清治 “野球のデータの統計的分析” 日本統計学会誌2007年3月 pp91~115

[6] 石岡文生、栗原考次、文勝浩 主成分分析を用いた空間データのホットスポットの検出について(日本計算機統計学会シンポジウム論文集 (19), pp39-42, 2005-10-20)

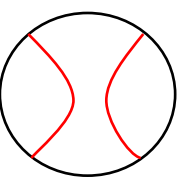


Appendix



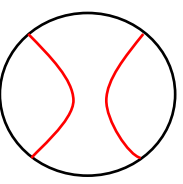
言葉の定義1

- 安打: 野球で、打者が守備側の失策なしに一塁またはそれ以上の塁に進むことができる打球。また、打者がそのような打球を飛ばすこと
- 打点: 野球で、打者が安打・犠打・四死球などで走者をホーム-インさせてあげた得点の数。
- 盗塁: 野球で、出塁した走者が守備側の隙をねらって、次の塁へ進むこと
- 四球: 野球で、打者が四つのボールを得ること。打者は一塁へ進むことができる
- 死球: 野球で、投球が打者の体または着衣に触れること。打者は一塁へ進塁することができる



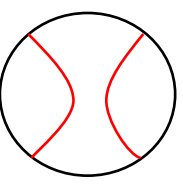
言葉の定義2

- 三振：野球で、打者がストライクを三回とられてアウトになること
- 併殺打：野球で、ダブルプレーをとられてしまう打撃
- ダブルプレー：野球で、2人の走者、あるいは1人の走者と打者を、連続したプレーによって合わせてアウトにすること
- 得点：走者が正規に一塁、二塁、三塁、本塁の順に進塁し、本塁に触れるときに記録される点数のことである。

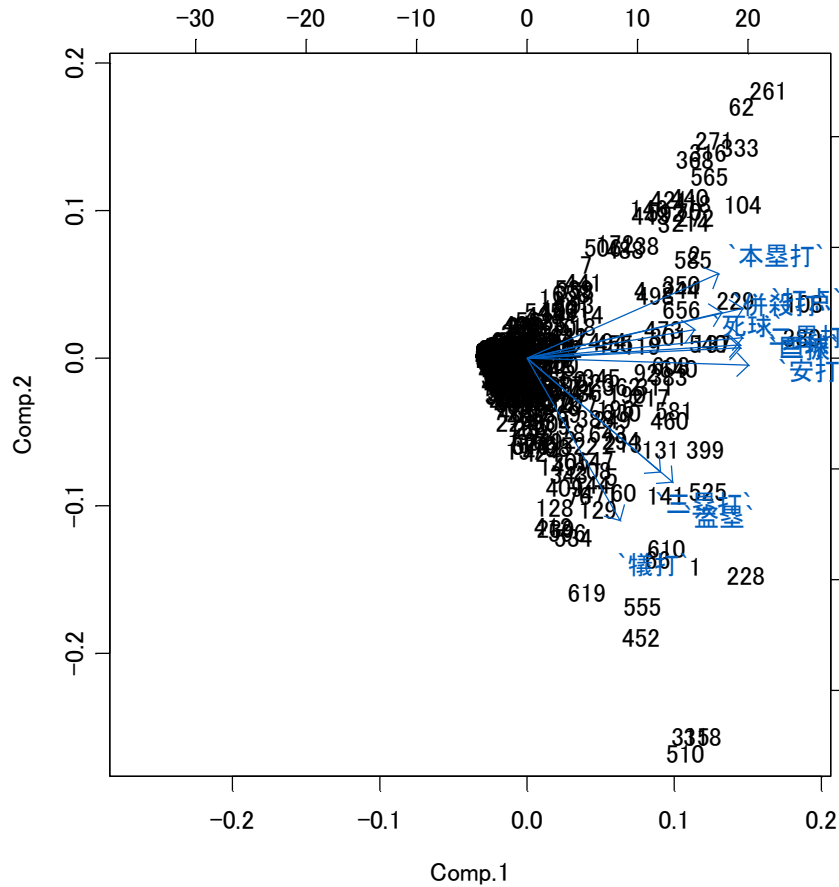


言葉の定義3

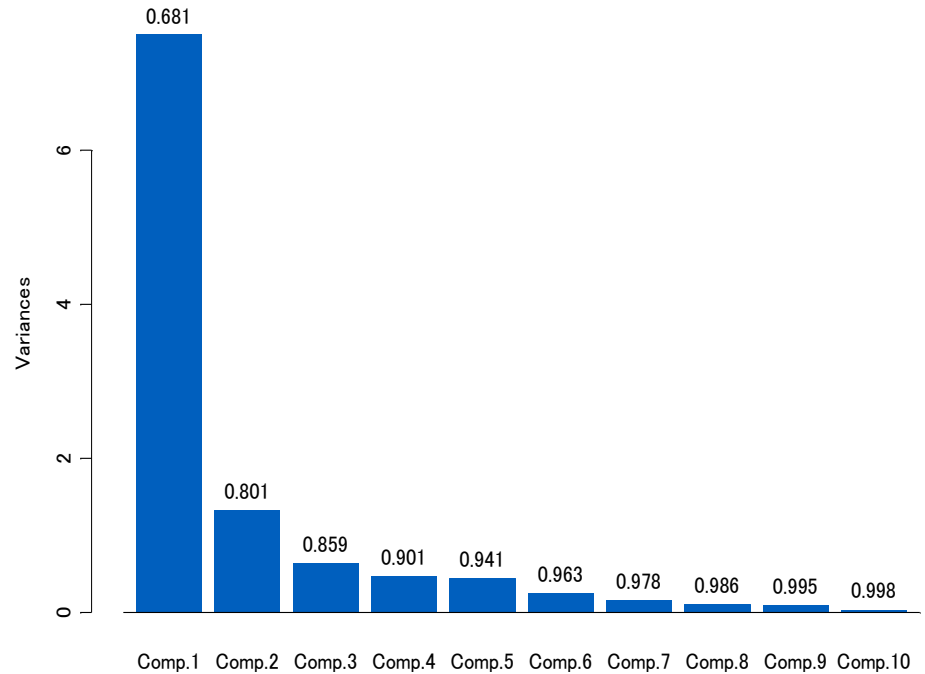
- 犠飛: 野球で、打者はアウトになるが、走者を進塁させたフライ。
- 塁打: 野球で、打った安打のうち、単打を1、二塁打を2、三塁打を3、本塁打を4として集計した数。
- 故意四球: 野球において投手が打者に対して、意図的に四球を与える行為につく記録である。

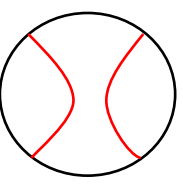


主成分分析の結果 (S-plus)



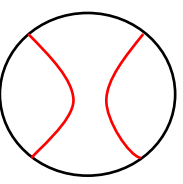
Relative Importance of Principal Components





主成分分析の結果 (S-plus)

- `princomp(x = ~ `安打` + `二塁打` + `三塁打` + `本塁打` + `打点` + `盗塁` + `犠打` + `四球` + `死球` + `三振` + `併殺打`, data = `野球1`, scores = TRUE, cor = FALSE, na.action = na.exclude)`
- `princomp(x = ~ `安打` + `二塁打` + `三塁打` + `本塁打` + `打点` + `盗塁` + `犠打` + `四球` + `死球` + `三振` + `併殺打`, data = `X2016野手データ...コピー`, scores = TRUE, cor = FALSE, na.action = na.exclude)`



主成分分析の結果 (VMS)

