

スポーツにおけるデータ分析

データからスポーツを探究する

GPSやセンサー、ドローン、IoT、データ蓄積技術、AI、データ分析アプリ、画像解析、などなど、近年の科学技術の発展はスポーツ界にデータ革命をもたらしました。情報通信技術やセンサー技術などの最新技術を活用し、プロスポーツやオリンピック競技を中心に様々なデータ収集と蓄積が急速に進んでいます。

Jリーグの試合で選手の走行距離が発表されたり、プロ野球の試合で投手の投げたボールの回転数が表示されたり、ゲーム中の選手の心拍数を測定することにより戦略的に選手を交代したり。このように、多くの競技では、それぞれの競技特性に応じて、プレー評価や戦略・戦術判断、パフォーマンス向上、選手育成、怪我の予防、およびチームマネジメントのためにデータ分析を活用しています。

本授業では、日本のプロスポーツにおけるケーススタディを通じて、データ活用の意義やその方法を理解することを目指しています。授業での気づきや学びを体育や部活動、さらにはスポーツシーン以外で実践することで、従来の方法を鵜呑みにせず、自分たちにあった練習方法、戦略・戦術等を考えられる力を養っていきましょう。

1. データから野球を考える

Challenge 01

セ・パリーグの主要投手のスタッツの中から2つを選び、例に従いそれぞれを横軸・縦軸にとったグラフを作成してください。次に、各投手をそれぞれのスタッツの対応する位置にマッピングしてみましょう。「この投手は○は高いが、○は低い」のように各投手の特徴を視覚的に把握することができます。いくつかマッピングし、特徴的な投手のプロフィールをグループで纏めてみましょう。

2020年セ・リーグ 主要投手の各種スタッツ

選手名	チーム名	防御率	奪三振率	与四球率	被本塁打率	ゴロ率	ストレート 投球割合	ストレート 平均球速	対右打者 被打率	対左打者 被打率
大野 雄大	中日	1.82	8.96	1.39	0.79	51%	51%	146.1	.189	.232
森下 暢仁	広島	1.91	9.10	2.35	0.44	50%	50%	147.7	.238	.222
菅野 智之	巨人	1.97	8.58	1.64	0.52	41%	34%	149.0	.177	.213
西 勇輝	阪神	2.26	7.01	1.71	0.91	53%	5%	141.9	.226	.198
九里 亜蓮	広島	2.96	7.30	3.03	0.76	54%	27%	143.3	.242	.235
青柳 晃洋	阪神	3.36	6.56	3.28	0.30	63%	14%	142.9	.193	.288

2020年パ・リーグ 主要投手の各種スタッツ

選手名	チーム名	防御率	奪三振率	与四球率	被本塁打率	ゴロ率	ストレート 投球割合	ストレート 平均球速	対右打者 被打率	対左打者 被打率
千賀 滉大	ソフトバンク	2.16	11.08	4.24	0.30	50%	46%	153.4	.187	.223
山本 由伸	オリックス	2.20	10.59	2.63	0.43	53%	34%	151.4	.172	.192
有原 航平	日本ハム	3.46	7.19	2.04	0.75	47%	28%	148.4	.233	.260
涌井 秀章	楽天	3.60	7.62	2.63	1.18	41%	44%	143.8	.230	.230
高橋 光成	西武	3.74	7.48	3.29	0.67	55%	42%	147.2	.238	.220
美馬 学	ロッテ	3.95	6.44	1.83	0.66	48%	21%	144.6	.213	.313
田嶋 大樹	オリックス	4.05	6.55	3.09	1.03	42%	47%	141.7	.241	.206
石川 歩	ロッテ	4.25	5.20	1.76	1.28	51%	30%	143.9	.273	.268

スタッツの解説

防御率

投手が9イニングあたりにどのくらい点を取られたかを表すスタッツ。
 防御率が**低いほど点を取られにくい投手**、防御率が**高いほど点を取られやすい投手**と言える。
 【計算式: 自責点×9÷投球回】

奪三振率

投手が9イニングあたりにどのくらい三振を奪ったかを表すスタッツ。
 奪三振率が**低いほど三振を奪うのが苦手な投手**、奪三振率が**高いほど三振を奪うのが得意な投手**と言える。
 【計算式: 奪三振×9÷投球回】

与四球率

投手が9イニングあたりにどのくらい四球を与えたかを表すスタッツ。
 与四球率が**低いほどコントロールが良い投手**、与四球率が**高いほどコントロールが悪い投手**と言える。
 【計算式: 与四球×9÷投球回】

被本塁打率

投手が9イニングあたりにどのくらい本塁打を打たれたかを表すスタッツ。
 被本塁打率が**低いほど本塁打を打たれにくい投手**、被本塁打率が**高いほど本塁打を打たれやすい投手**と言える。
 【計算式: 被本塁打×9÷投球回】

ゴロ率

投手が打たれた打球のうち、ゴロの打球がどのくらいあったかを表すスタッツ。
 ゴロ率が**低いほど打球がフライになりやすい投手**、ゴロ率が**高いほど打球がゴロになりやすい投手**と言える。
 【計算式: ゴロ打球数÷全打球数】

ストレート 投球割合

投手の全投球のうち、球種がストレートだった割合を表すスタッツ。
 ストレート投球割合が**低いほど変化球を多く使って勝負する投手**、ストレート投球割合が**高いほどストレートを多く使って勝負する投手**と言える。
 【計算式: ストレート投球数÷全投球数】

ストレート 平均球速

投手がストレートを投げた時のボールの平均速度を表すスタッツ。単位はkm/h。
 一般的に、ストレート平均球速が**速い投手ほど打者に打たれにくい**。
 【計算式: ストレート球速の合計÷ストレート投球数】

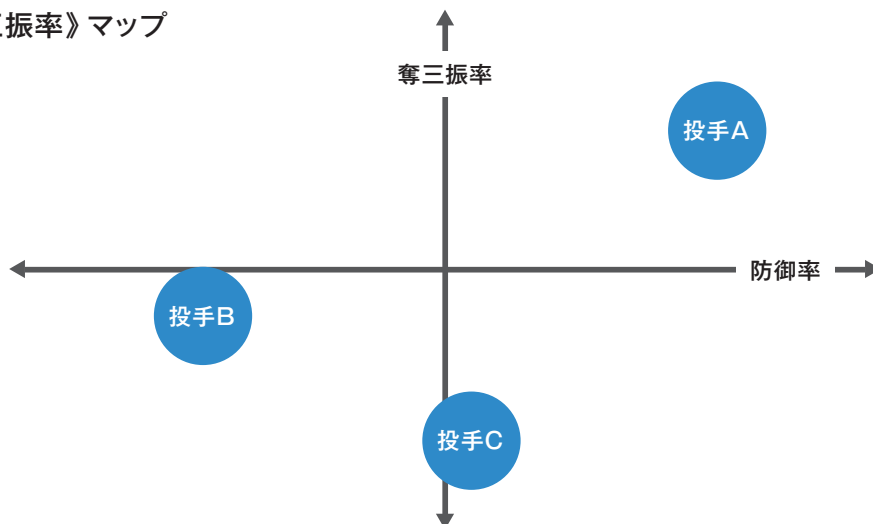
対右打者 被打率

投手が右打者と対戦した時にどのくらい安打を打たれたかを表すスタッツ。
 対右打者被打率が**低いほど右打者が得意な投手**、対右打者被打率が**高いほど右打者が苦手な投手**と言える。
 【計算式: 対右打者被安打÷対右打者打数】

対左打者 被打率

投手が左打者と対戦した時にどのくらい安打を打たれたかを表すスタッツ。
 対左打者被打率が**低いほど左打者が得意な投手**、対左打者被打率が**高いほど左打者が苦手な投手**と言える。
 【計算式: 対左打者被安打÷対左打者打数】

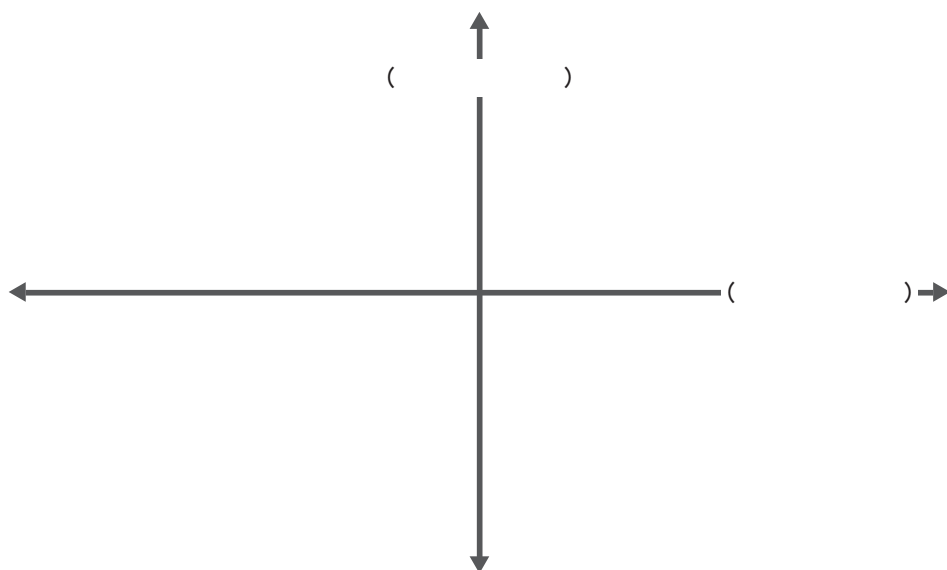
例) 《防御率×奪三振率》マップ



分析・考察

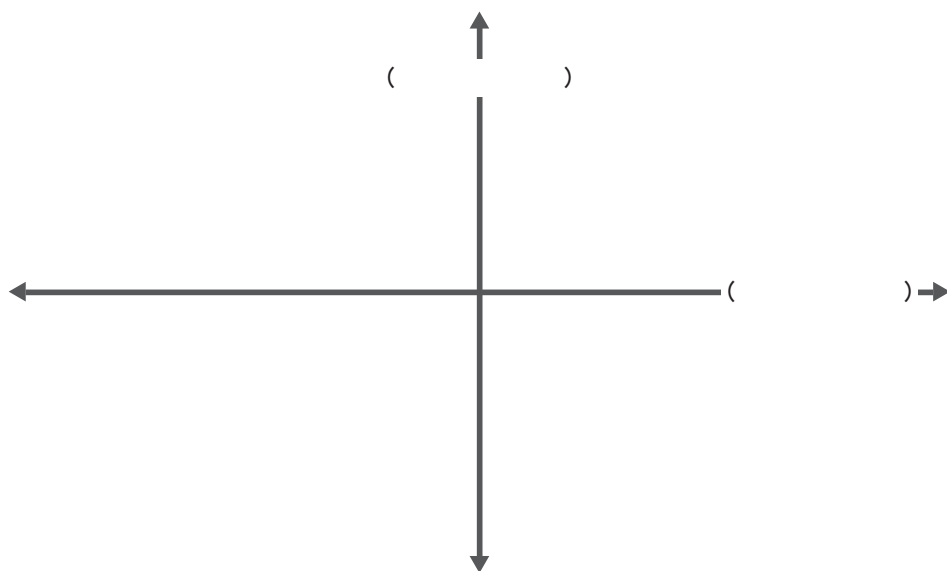
A 投手は奪三振率が高い割には防御率が高い（三振を多くとれる一方で失点が多い）
 →四球が多くてランナーをだしてしまうのかな…？ 球速は速いけど本塁打を打たれやすいのかな…？
 →与四球率はどんなだろう？ ストレート平均球速はどんなだろう？ 被本塁打率はどんなだろう？
 →（他のスタッツをクロスさせると、A 投手のプロフィールがもっと分かるかも…）

《 () × () 》マップ



分析・考察

《 () × () 》マップ



分析・考察

2. データからチームの課題を探る（野球編）

Challenge 02

2020シーズンのセ・パ両リーグの得失点スタッツをもとに、各リーグ最下位であったヤクルトとオリックスの課題（敗因）を見つけてみましょう。

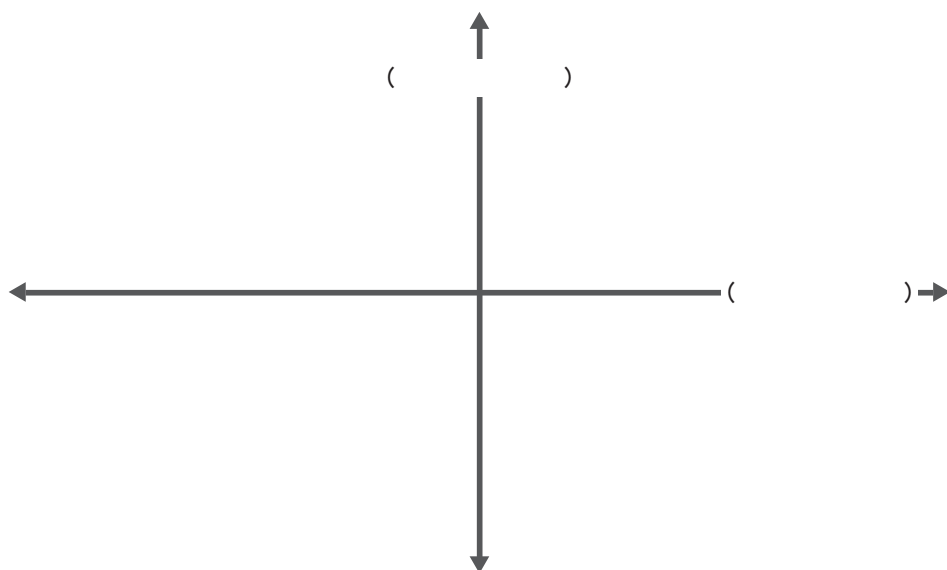
2020年セ・リーグチーム別得失点スタッツ

順位	チーム名	勝利	敗北	引分	勝率	得点	失点
1	巨人	67	45	8	.598	532	421
2	阪神	60	53	7	.531	494	460
3	中日	60	55	5	.522	429	489
4	DeNA	56	58	6	.491	516	474
5	広島	52	56	12	.481	523	529
6	ヤクルト	41	69	10	.373	468	589

2020年パ・リーグチーム別得失点スタッツ

順位	チーム名	勝利	敗北	引分	勝率	得点	失点
1	ソフトバンク	73	42	5	.635	531	389
2	ロッテ	60	57	3	.513	461	479
3	西武	58	58	4	.500	479	543
4	楽天	55	57	8	.491	557	522
5	日本ハム	53	62	5	.461	493	528
6	オリックス	45	68	7	.398	442	502

《 () × () 》マップ



分析・考察

特徴的な投手プロフィールのまとめ

Challenge 03

Challenge 02でオリックスの課題（敗因）がわかりましたが、さらにその原因がどこにあるのかデータを使いながら探してみましょう。

2020年パ・リーグ:チーム別攻撃スタッツ

チーム名	得点	打率	出塁率	長打率	ISO	本塁打率	三振率	四球率	盗塁成功率
楽天	557	.258	.341	.401	.143	.028	19.8%	10.2%	69.1%
ソフトバンク	531	.249	.321	.401	.152	.032	19.0%	8.9%	75.6%
日本ハム	493	.249	.330	.364	.115	.023	19.3%	10.1%	77.7%
西武	479	.238	.315	.374	.136	.027	21.5%	9.3%	68.0%
ロッテ	461	.235	.329	.355	.121	.023	21.2%	10.9%	73.1%
オリックス	442	.247	.316	.367	.120	.023	18.2%	8.1%	70.9%

打率＝安打÷打数
 出塁率＝（安打＋四球＋死球）
 ÷（打数＋四球＋死球＋犠飛）
 長打率＝塁打÷打数
 ISO＝長打率－打率
 本塁打率＝本塁打÷打数
 三振率＝三振÷打席
 四球率＝四球÷打席
 盗塁成功率＝盗塁÷（盗塁＋盗塁刺）

2020年パ・リーグ:チーム別投手・守備スタッツ

チーム名	失点	被打率	奪三振率	与四球率	被本塁打率	DER	守備率	盗塁阻止率
ソフトバンク	389	.219	8.74	3.86	0.70	.710	.987	25.6%
ロッテ	479	.252	7.12	3.26	0.94	.700	.988	25.2%
オリックス	502	.246	8.00	3.91	0.91	.693	.986	22.8%
楽天	522	.251	7.39	3.39	0.94	.696	.986	23.5%
日本ハム	528	.252	7.94	3.60	0.84	.680	.983	23.1%
西武	543	.256	6.55	4.09	0.91	.695	.985	30.1%

被打率＝被安打÷打数
 奪三振率＝奪三振×9÷投球回
 与四球率＝与四球×9÷投球回
 被本塁打率＝被本塁打×9÷投球回
 DER＝（打席－被安打－与四球－与死球－奪三振－失策）
 ÷（打席－被本塁打－与四球－与死球－奪三振）
 守備率＝（刺殺＋補殺）÷（刺殺＋補殺＋失策）
 盗塁阻止率＝盗塁刺÷（許盗塁＋盗塁刺）

Challenge 04

Challenge 02でヤクルトの課題（敗因）がわかりましたが、さらにその原因がどこにあるのかデータを使いながら探してみましょう。

2020年セ・リーグ:チーム別攻撃スタッツ

チーム名	得点	打率	出塁率	長打率	ISO	本塁打率	三振率	四球率	盗塁成功率
巨人	532	.255	.328	.410	.155	.034	21.6%	9.4%	70.8%
広島	523	.262	.331	.402	.140	.027	19.3%	8.4%	65.3%
DeNA	516	.266	.324	.415	.149	.033	19.3%	7.5%	67.4%
阪神	494	.246	.321	.381	.135	.028	21.3%	8.9%	70.8%
ヤクルト	468	.242	.321	.375	.133	.029	20.5%	9.5%	63.8%
中日	429	.252	.317	.360	.107	.018	18.3%	7.8%	60.0%

打率＝安打÷打数
 出塁率＝（安打＋四球＋死球）
 ÷（打数＋四球＋死球＋犠飛）
 長打率＝塁打÷打数
 ISO＝長打率－打率
 本塁打率＝本塁打÷打数
 三振率＝三振÷打席
 四球率＝四球÷打席
 盗塁成功率＝盗塁÷（盗塁＋盗塁刺）

2020年セ・リーグ:チーム別投手・守備スタッツ

チーム名	失点	被打率	奪三振率	与四球率	被本塁打率	DER	守備率	盗塁阻止率
巨人	421	.230	7.52	3.31	0.97	.721	.990	34.2%
阪神	460	.242	7.40	3.17	0.80	.693	.982	28.8%
DeNA	474	.255	8.03	3.38	0.84	.682	.988	33.3%
中日	489	.252	8.08	2.90	0.99	.687	.989	40.2%
広島	529	.262	7.69	3.61	0.95	.675	.984	31.9%
ヤクルト	589	.281	7.48	3.42	1.18	.661	.986	19.2%

被打率＝被安打÷打数
 奪三振率＝奪三振×9÷投球回
 与四球率＝与四球×9÷投球回
 被本塁打率＝被本塁打×9÷投球回
 DER＝（打席－被安打－与四球－与死球－奪三振－失策）
 ÷（打席－被本塁打－与四球－与死球－奪三振）
 守備率＝（刺殺＋補殺）÷（刺殺＋補殺＋失策）
 盗塁阻止率＝盗塁刺÷（許盗塁＋盗塁刺）