

Using Clustering to Find Pitch Subtypes and Effective Pairings

Gregory Dvorocsik, Eno Sarris, and Joseph Camp, PhD

INTRODUCTION

Using Statcast data, it is now possible to compare individual pitches across baseball based on characteristics like movement, velocity, and spin rate that become obvious and meaningful even in a single outing. Various research has used those physical characteristics to define optimal pitches.

However, to be effective even an elite pitch must be mixed with less optimal ones, especially for starting pitchers. Therefore, it is imperative to study the interactions between pitches to fully understand the best shape a particular pitch should have and whether pairing a pitch with others improves or decreases effectiveness. Here, we define pairing as when both pitches are present in a pitcher's arsenal.

This work has potential impact on pitch design, player development, and scouting. For the former, teams could focus efforts on teaching young pitchers new subtypes that have specific shapes according to the characteristics of the best pitches already thrown by that pitcher. For the latter, with very little in-game data, teams could seek to add pitchers that already possess effective pairings or avoid pitchers with ineffective pairings.

BACKGROUND

Pitch design has advanced to the point that tinkering in a lab with high-speed, high-resolution cameras is commonplace to potentially modify the ideal shape of a given pitch. These systems capture the way in which a pitch leaves the fingers, having potential impact on the spin axis, spin rate, velocity, and resulting horizontal and vertical movement. With such fine precision, the traditional classes of pitches might be called into question since sub-classes of pitch types could be forming based on a number of minor alterations to the grip, release point, and other pitching mechanics. We could potentially rely on existing works that have studied the following:

1. classifying pitches in an unsupervised manner pre- and post-Statcast¹
2. the characteristics of an effective changeup²
3. the optimal distribution of pitches³
4. a survey of publicly available data and existing methods to evaluate pitch types and performance⁴

However, these works have not evaluated how one pitch type or subtype can affect another when a pitcher throws multiple pitches. We seek to understand whether a particular pitch—with its physical characteristics of spin rate, velocity, and break—can have better or worse outcomes when a second pitch of different physical characteristics is paired with it.

In this paper, we use clustering within each of the most common pitch classifications as determined by Statcast to find subtypes of pitches and evaluate their performance when paired with all other subtypes to determine if performance can increase or decrease by these pairings. Looking at the 9 most common MLB pitch types for left-handed and right-handed pitchers across all pitches thrown in the 2016 and 2017 seasons, we find the total number of subtypes that result to be 30 for both types of pitchers. (Levels of subtypes differed between left-handers and right-handers.) We consider four performance metrics:

- swinging strike percentage (whiff rate)
- exit velocity (how hard the ball was hit)
- percentage of time the launch angle is below 0 percent (ground ball rate)
- percentage of time the launch angle is above 40 percent (pop-up rate)

We created a reference for each of these subtypes by studying the average performance of each of these 30 pitch subtypes for each handedness. We found that even within the same Statcast pitch types, there are

sizable differences in effectiveness for these four performance metrics across pitch subtypes.

Then, we compared this reference performance against the performance of that same reference subtype for all occurrences where a pitcher throws that second subtype and created lists ordered according to the change in each performance metric. Hence, we found the highest and lowest performing subtype pair for whiff, pop-up, and ground ball rates and exit velocity, calling these the best and worst subtype pairings for each reference subtype and performance metric. For example, looking at CH1 of LHP in the top left corner of Table 2 and Table 3, the highest positive difference in swinging strike rate for CH1 was when FC3 was paired with it, and the greatest negative difference in swinging strike rate for CH1 was when SI2 was paired with it. After repeating this for all reference subtypes and all possible subtype pairs across all performance metrics, we found that each of these best and worst pairing subtypes are typically distinct across these performance metrics for a given reference subtype. When we consider the average gains across each of these metrics, we see an improvement of 1.6 percent swinging strike percentage, 3.8 percent pop-up rate, 4.2 percent ground ball rate, and a reduction of 1.2 MPH on exit velocity. Conversely, by performing the same process for finding the worst pairing subtypes, we see reductions in the swinging strike rate by 1.9, pop-up rate by 2.9, ground ball rate by 4.8, and increase the exit velocity by 1.2 MPH.

PITCH SUBTYPE CLASSIFICATION USING STATCAST-DRIVEN K-MEANS CLUSTERING

In order to immediately broadcast pitch type across various platforms from in-stadium scoreboards to the MLB AtBat application, Major League Baseball classifies each pitch based on Statcast metrics. While we cannot know the exact algorithms used to perform the classification, based on per pitch data from 2016 and 2017, we know there are 9 predominant pitch types for both right-handed and left-handed pitchers:

- four-seam fastball (FF)
- two-seam fastball (FT)
- cutter (FC)
- splitter (FS)
- sinker (SI)
- curveball (CU)
- knuckle curve (KC)
- slider (SL)
- changeup (CH)

Since these classifications are broad in nature, there are differences in horizontal and vertical movement, velocity, and spin rate, even within the same class of pitch. Movement of pitches is relative to a gyroball, which is a ball spinning in a spiral shape (like a football). A gyroball-shaped pitch is considered the theoretical zero/zero from which other pitches are defined. Unlike a perfect gyroball, almost all pitches exhibit some form of magnus effect—the forces that deflect the ball in a particular direction based on the velocity, spin, and spin axis. To understand the degree to which these four physical characteristics differ within each class, we use clustering within a four-dimensional space as represented by these four physical parameters:

Horizontal Movement: From the catcher's viewpoint and with respect to the movement of a gyroball, a negative value would move toward a right-handed hitter, whereas a positive value would move toward a left-handed hitter.

Vertical Movement: From the catcher's viewpoint and as compared to the movement of a gyroball, a negative value would move downward, whereas a positive value would rise.

Velocity: The miles per hour that a pitch travels as measured out of the pitcher's hand.

Spin Rate: The revolutions per minute along the spin axis of a particular pitch.

We use K-means clustering to form the clusters, meaning we partition n observations into the k clusters where there exist an n of the number of pitches thrown in 2016 and 2017 from a particular handedness from a particular MLB pitch classification. The k is determined via the elbow method where with each of the increasing k values, we evaluate the aggregate Within-Cluster-Sum-of-Squares (WCSS) error between all the data points and the k cluster centroids. In particular, over all k for a given MLB pitch type, when the WCSS value begins to flatten, an elbow is created in

LHP Type/ Subtype	Count	SwK (%)	Exit Velo (MPH)	LA>40° (%)	LA<0° (GB %)	Velo (MPH)	Spin (RPM)	Xbrk (in.)	Ybrk (in.)	RHP Type/ Subtype	Count	SwK (%)	Exit Velo (MPH)	LA>40° (%)	LA<0° (GB %)	Velo (MPH)	Spin (RPM)	Xbrk (in.)	Ybrk (in.)
CH	45948	15.7	81.3	16.0	37.1	83.1	1811	7.0	4.8	CH	97836	15.0	82.1	15.0	37.8	84.3	1741	-6.3	5.5
CH 1	18837	15.7	82.0	17.1	36.3	84.1	1961	6.3	6.3	CH 1	28814	14.1	83.5	14.1	38.4	86.6	1912	-5.6	5.6
CH 2	13286	15.9	82.2	13.9	39.9	81.2	1529	6.5	5.9	CH 2	27372	16.0	82.8	12.1	42.9	83.4	1480	-5.7	5.3
CH 3	13825	15.5	79.5	16.6	35.5	83.5	1878	8.5	6.0	CH 3	16232	15.3	81.2	21.4	29.1	79.7	1901	-6.0	6.0
CH 4	28233	12.5	81.5	13.0	38.8	76.7	2440	-3.0	-2.7	CH 4	25418	14.7	80.5	14.9	37.7	85.4	1725	-7.8	5.4
CU	6561	12.1	79.2	10.9	42.8	75.7	2551	-4.2	-3.7	CU	87674	12.2	81.8	13.9	36.0	78.1	2498	4.0	-4.7
CU 1	10644	13.2	82.4	16.0	33.1	79.0	2259	-3.1	-3.0	CU 1	35888	10.7	82.5	12.3	38.2	76.1	2589	3.4	-5.0
CU 2	11028	12.0	82.1	11.2	42.4	75.1	2550	-2.4	-3.7	CU 2	31387	13.9	82.2	16.5	32.5	80.6	2336	3.9	-4.3
CU 3	15322	10.6	82.1	16.7	32.8	87.4	2193	0.6	5.6	CU 3	20399	12.3	80.1	12.8	37.4	77.8	2589	5.1	-4.9
FC	5524	11.8	81.8	16.0	33.0	86.1	2278	0.8	6.4	FC	50020	11.4	81.9	19.6	29.6	88.7	2363	-0.1	7.3
FC 1	4876	9.4	82.2	15.9	33.5	86.8	1995	0.4	6.6	FC 1	17663	10.8	81.3	22.7	27.2	91.7	2441	0.0	7.7
FC 2	4922	10.6	82.3	18.3	31.8	89.4	2292	0.7	6.9	FC 2	12402	9.5	82.5	19.5	30.6	87.3	2091	0.0	7.5
FC 3	13063	7.7	83.1	25.4	21.7	92.2	2234	5.4	8.6	FC 3	19955	13.1	82.1	17.1	31.1	86.9	2462	-0.1	7.0
FF	34481	7.6	81.9	24.5	21.9	92.6	2206	6.4	10.2	FF	376138	7.9	83.6	26.1	20.6	93.1	2269	-4.7	10.6
FF 1	41537	9.9	84.5	21.5	26.6	90.4	2110	4.9	10.2	FF 1	103554	8.0	82.1	25.2	20.8	93.9	2266	-5.7	10.5
FF 2	54045	5.1	82.9	29.1	17.6	93.2	2348	5.1	10.6	FF 2	124287	5.8	84.6	24.2	23.3	91.1	2145	-4.3	10.5
FF 3	2679	18.7	81.8	16.4	38.7	83.4	1562	8.3	3.8	FF 3	148297	9.4	84.0	28.7	18.1	94.4	2376	-4.3	10.8
FS	503	20.1	83.9	13.3	41.0	84.6	1474	7.5	4.6	FS	16782	17.2	82.5	13.0	40.6	84.7	1516	-5.3	4.3
FS 1	863	20.3	83.9	11.3	42.1	81.3	1255	7.7	4.4	FS 1	5310	18.1	82.5	14.6	36.9	82.2	1241	-5.1	4.1
FS 2	719	13.9	78.9	30.0	29.4	84.2	2154	8.5	5.4	FS 2	7366	16.2	83.7	11.8	44.0	86.3	1693	-4.7	4.5
FS 3	594	21.2	79.5	11.7	42.2	84.5	1368	9.6	4.5	FS 3	4106	18.0	80.6	13.2	39.6	85.3	1554	-6.8	4.2
FT	55588	6.1	84.1	15.6	35.4	91.5	2134	8.1	5.9	FT	144139	5.7	84.5	15.6	35.1	92.4	2170	-7.9	7.7
FT 1	23008	7.0	85.2	18.4	32.4	92.8	2206	7.3	7.6	FT 1	46996	5.7	82.2	16.0	33.8	92.6	2155	-9.5	7.5
FT 2	15669	5.0	85.0	12.6	41.5	89.4	2030	7.5	7.2	FT 2	47317	6.5	85.7	18.4	31.5	93.9	2298	-7.2	7.9
FT 3	16911	5.8	82.0	14.8	34.1	91.7	2132	9.7	7.3	FT 3	49826	4.9	85.7	12.5	39.9	90.9	2062	-7.0	7.6
KC	8318	12.6	82.7	8.6	41.9	79.3	2229	-2.7	-3.3	KC	22914	13.3	82.5	10.6	43.5	80.7	2520	4.2	-6.0
KC 1	2172	12.1	83.9	8.3	40.8	78.8	2474	-2.2	-3.9	KC 1	7884	9.6	84.4	9.7	44.3	78.2	2424	3.8	-6.2
KC 2	1571	10.4	81.1	5.6	44.4	77.7	2350	-3.6	-4.4	KC 2	6206	17.6	82.7	10.8	45.4	83.4	2732	3.7	-5.9
KC 3	1994	12.2	84.2	4.6	56.9	79.4	2137	-2.3	-4.5	KC 3	6056	13.4	80.1	10.0	43.7	80.9	2615	5.1	-6.2
KC 4	2581	14.7	81.9	13.3	31.3	80.7	2022	-2.9	-3.4	KC 4	2768	14.0	82.3	13.6	37.3	81.0	2113	4.4	-5.5
SI	23387	6.8	83.7	17.1	35.9	90.8	2107	7.9	4.5	SI	59971	5.5	84.8	13.2	38.5	91.6	2124	-7.8	6.1
SI 1	13200	6.8	84.6	21.1	30.0	91.7	2172	7.4	6.3	SI 1	26596	5.9	85.6	14.8	35.9	93.1	2230	-7.3	6.3
SI 2	4846	6.3	84.6	8.9	47.3	87.5	1953	7.3	5.6	SI 2	12880	5.5	82.4	14.1	38.2	91.8	2126	-9.9	6.0
SI 3	5341	7.2	81.1	14.9	40.0	91.4	2083	9.6	5.8	SI 3	20495	5.1	85.6	10.6	42.0	89.5	1984	-7.3	5.9
SL	44923	15.9	81.8	17.0	33.9	83.1	2275	-1.4	1.4	SL	142297	15.8	81.6	18.4	32.4	85.0	2345	2.0	2.3
SL 1	10902	16.2	79.6	18.2	32.3	80.9	2307	-2.1	1.4	SL 1	32906	15.5	82.6	18.6	30.3	82.2	2457	1.4	2.0
SL 2	20593	15.9	82.3	17.0	35.3	85.9	2326	-1.4	2.0	SL 2	71972	15.5	81.6	18.2	33.3	87.1	2338	2.0	2.6
SL 3	10881	15.7	82.6	15.1	33.5	80.4	2385	-0.7	1.3	SL 3	6199	15.3	83.0	17.9	33.4	84.2	1278	1.9	2.4
SL 4	2547	15.7	83.0	18.2	30.5	81.0	1252	-1.5	1.9	SL 4	31220	17.0	80.2	18.8	32.3	83.4	2457	2.6	2.0

Table 1. LHP and RHP type/subtype performance metrics and physical characteristics.

the curve, signaling the reduction of error by increasing the number of clusters has lessened significantly and forming the appropriate number of clusters. To ensure that each physical trait does not dominate, we scale the smallest and largest value of each field to be in the same range of -1 to +1. Since the elbow method is known to suffer from some ambiguity in terms of the precise location of the bend, we have favored a lower number of clusters to avoid multiple versions of the same subtype.⁵ In other words, we use the inflection point where the slope first begins to flatten.

Interestingly, while each number of pitch subtypes is not the same per MLB Statcast pitch type, the total number of pitch subtypes ends up being 30 for both right-handed pitchers (RHP) and left-handed pitchers (LHP). We now consider varying levels of performance across the types and subtypes for each pitch. Since pitches could be successful in different manners, we choose different performance metrics evaluated when each pitch type or subtype is thrown:

1. swinging strike percentage, meaning the percentage of all pitches thrown of that kind that generate swings and misses,
2. average exit velocity, meaning the MPH off the bat when contact is made with that kind of pitch, and

3. extreme launch angles, which we define as the angle off the bat for all batted balls off of that kind of pitch that are either less than 0 degrees (ground ball) or greater than 40 degrees (pop fly).

We first consider the aggregate performance of all pitches thrown in each subtype regardless of what other types of pitches are combined to form a pitch arsenal. Table 1 captures all 9 of the predominant pitch types for left-handed pitchers and right-handed pitchers on the left and right side of the table, respectively. Below each of the 9 types are the subtypes as defined by the aforementioned k-means clustering and elbow method. For the type and subtypes, we have presented the count of the total number of times each pitch has been thrown over the two seasons, the performance metrics, and the four physical dimensions over which we based the clustering.

The performance of some subtypes can be vastly different, even for the same type of pitch. For example, RHP KC1 has a swinging strike percentage of only 9.6 percent as compared to an average of 13.3 and as high as 17.6 for KC2. The key distinction between KC1 and KC2 subtypes is that the lower-performing KC1 is 5.2 MPH slower but has a similar pitch shape in terms of horizontal and vertical break. To give a feel for another performance metric, we turn to the situation where the launch angle is above 40 degrees, producing

a routine fly ball. For LHP, a high-spin splitter (FS3) produces a fly ball at a 30 percent rate versus 11.3 percent for a low-spin splitter (FS2). This relationship is flipped for changeups, where there is an advantage to reduced spin for increasing the ground ball rate (launch angle below 0 degrees). This can be seen in the low-spin version of the changeup (CH2 for both) versus a changeup with a higher spin (CH3 for both) with the RHP version increasing the ground ball rate by 13.8 percent.

We can more easily see the physical characteristics of spin in revolutions per minute (RPM) and the horizontal and vertical break in inches in Fig. 1, where we have separated the spin and velocity into one graph and horizontal-vertical break into another graph. We observe that there is more horizontal-break diversity in subtype characteristics than vertical-break diversity. For example, while subtypes rarely span more than 1 or 2 vertical inches, they can have over 6 inches of difference in horizontal movement (e.g., RHP FT). There is far greater distinction in the spin and velocity of subtypes, as observed in Fig. 1. For example, the cluster centers for sliders can vary by more than 1200 RPMs and over 7 MPH for changeups.

When we focus on the shape of the spin-velocity figures by types, we can observe the two places that LHP and RHP had most dissimilar subtype clustering. First, LHP had a high-spin splitter that was not present in the RHP figure. Second, RHP had a changeup that had a relatively higher spin rate for the lower velocity of the pitch. Lastly, while we observe the same number of knuckle curves for both pitcher types, the shape is far more distinct for RHP, where there is a high-spin, high-speed version that is not present with LHP.

EVALUATING THE EFFECTIVENESS OF PITCH SUBTYPE PAIRINGS

Our goal in this section is to quantify the effect of a single pitch subtype when the pitcher pairs that subtype with another subtype. In other words, when a pitcher has both subtypes in his arsenal, we say that the two subtypes have a pairing. To do so, we consider any time these reference subtypes are paired with each of the other subtypes, meaning the pitcher throws both subtypes. We evaluate all combinations of pitch pairings according to the same four performance metrics, as introduced in Section 2. In other words, we compare the aggregate performance of a pitch subtype

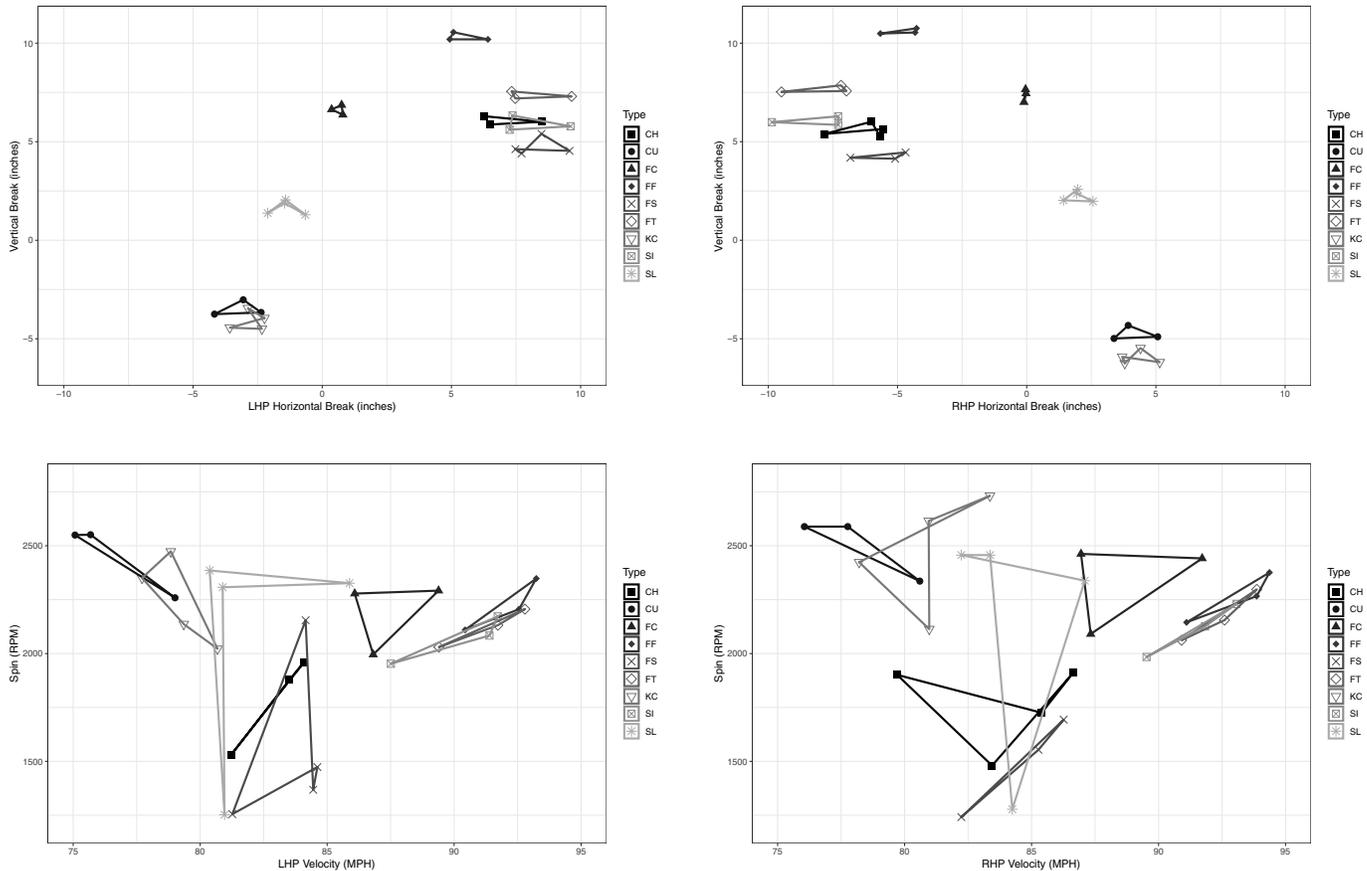


Figure 1: Four physical dimensions for pitch subtype clustering separated into spin-velocity (top) and horizontal-vertical break (bottom) for left-handed (left) and right-handed (right) pitchers.

(shown in Table 1) against the performance of that same pitch when paired with each of the 29 other subtypes to determine the most extreme gains and losses in performance. To do so, we create ordered lists for each of the performance metrics for each reference subtype when paired with all other subtypes, using a cutoff of at least 100 pitches thrown. Almost all of our subtypes have a sample of at least 200 thrown, where most of our pitch type statistics should be considered statistically “stable.”⁶ We left a few smaller-sample pitch types (four in total) in the study in order to improve the breadth of pitch types included. Then, we identify the best subtype pair (Table 2) and worst subtype pair (Table 3) and show the difference (Δ) in performance from the reference of all occurrences of that subtype being thrown, presented in Table 1.

First, we generally observe the difference in physical characteristics that have extreme gains along each of the performance metrics. Across both RHP and LHP changes in the vertical movement between the reference subtype and paired subtype led to improvements in popup percentage but at the cost of reductions in ground ball percentage. For improvements in exit velocity for both LHP and RHP, a large difference in

horizontal break between the reference and paired subtype is very beneficial. We observe that for increases in swinging strike percentage changes in velocity from the reference subtype and the paired subtype is a primary driver for RHP. For LHP, we found that changes in vertical break between the reference subtype and the paired subtype also played a significant role in swinging strike improvements.

We now discuss some noteworthy observations from Table 2 and Table 3, going from top to bottom. We find that there is an interesting trend for the best changeup pairings based on different performance metrics. We find that the greatest swinging strike rate happens when there is a high level of distinction in at least one physical factor from the reference of a changeup, but distinct by handedness: LHP experience the best swinging strike pairings with subtypes that have the greatest horizontal separation from the changeup, whereas the greatest change in velocity is the key for RHP. Conversely, the biggest changes in exit velocity occur for both handedness when pitches are of similar break to changeups, inducing weak contact. What is striking with the poor pairings for changeups, especially for RHP, the FS that is most

LHP Subtype	Δ Swk (%)	Δ Exit Velo (MPH)	Δ LA>40° (Pop-up %)	Δ LA<0° (GB %)	RHP Subtype	Δ Swk (%)	Δ Exit Velo (MPH)	Δ LA>40° (Pop-up %)	Δ LA<0° (GB %)								
CH1	FC3	1.3	FC2	-0.3	SI3	2.9	FC1	3.2	CH1	FC1	0.2	SI3	-0.5	SI1	1.8	SI2	2.1
CH2	KC4	1.9	SI3	-1.4	KC2	1.2	FS1	9.2	CH2	FS1	3.0	FS1	-1.6	FS1	2.6	SI2	3.8
CH3	FC2	1.8	FT2	-0.3	FF3	0.8	FC1	1.7	CH3	SI1	2.5	FS2	-2.3	FS1	14.6	KC2	4.6
CU1	FC2	0.7	SI3	-1.2	SI3	5.5	FF3	1.5	CH4	FF3	0.2	SI1	-0.7	CH3	0.8	SI3	1.2
CU2	SI3	1.2	SI3	-0.4	SL1	0.3	SI3	4.7	CU1	SI3	1.4	FS1	-0.7	FS1	2.4	KC2	12.1
CU3	SI3	4.2	SI3	-1.4	SI3	1.5	SL1	2.8	CU2	SI2	3.6	SI3	-1.6	FS1	3.9	KC4	10.7
FC1	KC3	2.8	SI3	-0.3	KC1	5.1	SI3	3.9	CU3	SI3	2.4	FS2	-0.9	FS1	4.4	FS2	4.9
FC2	SI2	1.9	FS1	-1.6	FS1	3.8	FS1	5.2	FC1	SL1	0.8	SL4	-0.5	FS3	7.4	SI2	1.3
FC3	SI3	1.4	SI3	-1.9	SI2	2.8	SL1	3.0	FC2	FT2	0.9	KC2	-1.8	FS3	1.9	FC3	1.3
FF1	FS2	0.8	FC3	-0.4	FS2	6.1	SI1	3.1	FC3	SL3	2.5	SL3	-0.4	KC1	2.1	SL3	1.4
FF2	SI3	1.0	SI3	-1.3	KC1	2.1	FS4	6.6	FF1	KC2	0.6	FS2	-0.4	FC1	1.1	SI2	2.1
FF3	FS2	1.6	KC2	-0.4	FS2	6.0	SI3	1.1	FF2	SI1	0.6	SI1	-0.5	FS1	2.6	SI2	4.1
FS1	FC3	2.5	SL4	-3.1	SL4	10.9	SL3	0.7	FF3	KC4	0.8	SI3	-0.3	FS2	1.6	SI2	3.5
FS2	CU3	0.4	SL3	-1.5	SL3	2.3	FS4	0.5	FS1	CH1	1.5	CH1	-1.6	KC3	4.0	SI2	8.1
FS3	FT1	0.4	CU3	-6.0	FT2	1.6	FS1	0.2	FS2	CH2	1.4	CH3	-0.7	CH4	2.7	SI2	6.8
FS4	CH2	2.8	CU3	-1.6	CU2	3.8	CH2	2.7	FS3	SI2	3.1	SI2	-1.4	CH3	3.4	SI1	7.9
FT1	FS1	0.7	KC1	-1.2	KC1	5.6	SI2	3.5	FT1	SI2	0.3	SI1	-0.2	KC2	2.0	FS2	4.3
FT2	CU2	0.4	KC3	-1.2	KC1	9.2	SI1	4.7	FT2	SI2	1.2	FS1	-1.4	SI2	6.1	KC3	1.6
FT3	SI3	3.2	SI3	-1.2	KC1	4.7	SL4	3.0	FT3	FS3	0.6	CH3	-0.2	KC1	0.7	FC1	1.9
KC1	SL3	2.4	CH3	-0.1	FT3	1.7	FC2	4.8	KC1	FS3	3.3	SI3	-2.4	FC3	3.0	CU2	3.3
KC2	SL2	3.3	FT3	-0.7	SL2	3.2	FC2	3.5	KC2	FC2	3.1	SL4	-1.3	SL4	2.8	FC3	4.0
KC3	SL2	0.5	SL2	-0.8	SL2	1.6	FT1	1.5	KC3	FC3	0.8	SI1	-2.2	SI1	2.2	CU2	5.4
KC4	KC2	1.6	KC2	-1.0	SI3	6.6	KC2	5.1	KC4	KC2	1.3	KC2	-0.9	FS3	4.8	CU2	5.4
SI1	CU2	0.6	KC2	-0.6	FC2	8.8	FT2	7.9	SI1	FS3	2.0	FT3	-1.7	FT3	3.1	KC1	3.5
SI2	FF1	1.5	FT1	-1.5	CU1	5.9	FC3	6.1	SI2	SL4	0.3	FT3	-0.7	FS1	3.9	KC1	4.0
SI3	FF1	0.4	FF3	-1.5	CU1	5.5	FC2	13.0	SI3	FT2	0.8	FF3	-0.6	FS1	3.6	KC4	4.5
SL1	FS2	1.2	FS4	-3.3	FS4	2.8	FS3	4.9	SL1	KC2	2.3	SI2	-1.8	KC4	2.8	KC2	7.3
SL2	SI2	0.9	KC1	-1.5	SI3	6.8	FC2	3.2	SL2	SI1	0.9	SI2	-0.8	FS2	3.3	FF3	0.1
SL3	FC2	0.7	FS4	-1.0	CU3	2.7	FC1	4.3	SL3	KC1	2.9	SI3	-2.2	KC1	3.9	FS3	4.6
SL4	FS1	1.2	CU3	-0.9	FS3	0.9	CU1	4.8	SL4	KC2	4.3	SI2	-0.8	FS1	4.0	KC2	2.8
AVG. GAIN		1.5		-1.3		4.1		4.0	AVG. GAIN		1.7		-1.1		3.5		4.3

Table 2. Best subtype pairings in terms of gain in each of the four performance metrics for the reference subtype for LHP (left) and RHP (right).

similar in break and velocity to the reference changeup subtype (see CH1, FS3, CH3, and FS1 in Table 1) dramatically reduces the swinging strike and ground ball rate outcomes of the changeup.

Sinkers (especially SI3) pair fairly universally well with curveballs, especially across performance metrics for LHP but even RHP. However, a lower-spin splitter (FS1) is the best match for all RHP curveball types for pop-up percentage and the more downward knuckle curve dramatically helps the CU to induce more ground balls. In fact, SI can pair both well and poorly with FS and driven by the velocity separation, since they have very similar movement and can have large spin differences with poor pairings.

For RHP cutters (FC), we see a compelling opposite trend. Namely, cutters play up (Table 2) when they are paired with pitches with more positive horizontal break and play down (Table 3) with pitches with more negative horizontal break. Conversely, RHP splitters play up when they are paired with pitches with more negative horizontal break and play down when paired with pitches with more positive horizontal break. LHP cutters play up with pitches with more negative horizontal break, such as sinkers for swinging strike rate

and exit velocity. When SL2 or SL3 is paired with the cutter, the differing horizontal and vertical break induces more ground balls. When SL1 or SL4 is matched with the curveball (SL1, CU2, SL4, and CU1) both losses occur of swinging strike rate going down and stronger contact being induced. The poor pairings seemingly result from differing vertical break between the pitch subtypes.

Both fastball (FF and FT) subtypes seem to be the least affected by the pairings in terms of swinging strike differences (1.6 Δ). Notice though that high-spin FF pairs better with low-spin KC (and better velocity separation) versus low-spin FF with high spin KC. Also, notice that the highest speed, highest spin class of FT pairs best with SI and worst with the slowest changeups. With lefties, SI2 or SI3 is paired well with FF2 or FF3 with respect to swinging strike rate. Four-seam fastballs pair poorly with splitters with regards to failing to induce ground balls (significant spin difference and vertical break change).

For RHP, knuckle curves can have very positive pairings with FS (similar velocity but very different vertical and horizontal movement) and FC (similar horizontal movement but substantial speed difference and vertical

LHP Subtype	ΔSwK (%)	ΔExit Velo (MPH)	ΔLA>40° (Pop-up %)	ΔLA<0° (GB %)	RHP Subtype	ΔSwK (%)	ΔExit Velo (MPH)	ΔLA>40° (Pop-up %)	ΔLA<0° (GB %)								
CH1	SI2	-2.1	SI1	0.5	FC1	-1.9	SI2	-2.2	CH1	FS3	-3.2	FS2	1.9	FC3	-1.2	FS3	-6.8
CH2	CU3	-0.6	FS1	2.7	SI2	-2.6	CU2	-1.6	CH2	KC3	-0.7	CU2	0.4	SI2	-1.7	FS2	-2.5
CH3	KC4	-1.0	KC4	1.3	KC4	-3.0	KC2	-2.0	CH3	FS1	-2.6	KC3	2.0	KC1	-4.4	FS1	-11.0
CU1	SL3	-0.4	SL3	1.0	FC1	-2.8	SL4	-4.0	CH4	KC4	-2.0	KC1	1.1	KC2	-2.8	CH3	-1.1
CU2	SL3	-0.6	FC1	0.9	SI1	-1.7	FC3	-1.0	CU1	FS1	-2.3	KC2	3.1	KC2	-4.5	FS1	-6.0
CU3	SL1	-2.2	FC1	0.8	FT1	-1.3	SI2	-5.4	CU2	FS1	-4.0	SL4	0.5	KC4	-7.2	FS1	-6.5
FC1	CU3	-1.2	SL3	1.0	SI3	-1.2	KC1	-6.1	CU3	SL4	-1.2	KC4	1.6	KC4	-1.6	SI2	-3.5
FC2	FS1	-2.5	FF3	0.2	FF3	-0.2	SI1	-3.3	FC1	CH3	-1.6	CH3	0.7	CH3	-1.0	FS3	-7.6
FC3	SL3	-2.3	KC3	1.6	SL1	-2.0	CU1	-1.5	FC2	SI3	-1.9	FS3	2.1	SI3	-1.7	FS3	-3.8
FF1	SI2	-0.9	FS1	0.7	SI2	-3.4	FS2	-6.5	FC3	FS3	-3.5	FT2	0.3	SI3	-1.5	KC2	-3.4
FF2	KC4	-0.7	KC1	2.9	FS4	-5.3	FS3	-2.6	FF1	SI2	-0.9	KC1	0.3	SI2	-3.0	FS2	-1.7
FF3	SI3	-2.7	FS4	1.9	SI3	-5.1	FS2	-3.1	FF2	KC3	-0.9	KC2	0.5	SI2	-3.8	FS1	-1.5
FS1	SL1	-2.2	CU3	1.4	CU3	-7.0	SL4	-4.6	FF3	CU2	-0.4	FS1	0.3	SI1	-1.8	CH3	-1.1
FS2	SL1	-1.8	FC2	0.3	CU2	-0.8	SL1	-1.5	FS1	SI2	-2.3	FC1	2.4	SI2	-4.6	KC3	-7.4
FS3	FF2	-0.1	SL4	1.0	CU3	-6.2	FT1	-1.8	FS2	FC2	-2.6	CU3	0.9	SI2	-2.3	CH4	-4.3
FS4	FT3	-2.0	CH1	2.2	CH2	-7.4	FT1	-2.7	FS3	CH3	-3.6	KC1	2.3	KC4	-3.2	CH3	-4.7
FT1	SI1	-1.7	CU3	0.5	SI2	-1.7	KC2	-6.7	FT1	SI3	-0.6	KC1	1.3	SI3	-1.9	KC2	-4.5
FT2	KC2	-1.3	FC3	0.5	SL4	-1.8	KC2	-11.8	FT2	CH3	-0.6	SI2	2.5	CH4	-0.9	SI2	-10.3
FT3	CU3	-1.1	FC2	0.5	SL4	-1.5	KC1	-10.7	FT3	SI3	-0.9	KC3	1.0	FS2	-1.8	KC1	-4.4
KC1	FT3	-1.6	FC2	1.2	FC2	-4.0	FT3	-5.8	KC1	SI2	-2.1	FC3	0.3	CU2	-5.7	CU1	-6.5
KC2	FT3	-1.3	KC4	0.4	KC4	-1.6	FT3	-5.3	KC2	SI2	-6.4	SI2	1.0	CH3	-1.2	CU3	-5.0
KC3	FT2	-0.4	FT1	0.7	FT1	-0.7	SL2	-1.5	KC3	SL1	-3.0	CU1	1.3	CU2	-3.1	SI3	-5.1
KC4	SL4	-4.1	SI2	1.8	FC3	-3.6	SI3	-10.9	KC4	SL3	-3.5	SL3	1.2	CU2	-4.7	SL3	-9.2
SI1	KC1	-2.0	FC1	0.8	FT2	-5.5	FC2	-6.8	SI1	KC1	-1.0	CH3	0.8	KC1	-1.3	FS1	-5.0
SI2	CU3	-1.3	FC2	2.5	FC3	-2.9	CU3	-10.8	SI2	FS3	-1.3	FS1	1.6	KC1	-2.8	FS1	-6.3
SI3	FC1	-1.2	SL4	0.5	FT2	-5.2	CU1	-9.9	SI3	FS1	-1.2	FS2	0.9	KC1	-1.7	FS1	-7.0
SL1	FC2	-2.4	FC2	1.2	FS3	-4.2	CU1	-3.7	SL1	FS1	-3.2	FC2	0.5	KC2	-5.8	CU1	-1.4
SL2	FC3	-4.0	CU3	0.5	FC2	-1.9	KC1	-3.2	SL2	KC1	-2.3	FS1	0.6	KC1	-0.9	SI2	-2.5
SL3	FS4	-3.9	CU3	0.8	FS4	-3.3	CU3	-4.1	SL3	CU3	-1.4	FT2	0.4	FC2	-2.9	CH1	-0.5
SL4	FC1	-2.0	FC1	1.2	CU1	-1.3	FC3	-4.3	SL4	CU3	-1.0	KC1	1.7	FC2	-0.6	FS1	-4.7
AVG. LOSS		-1.7		1.1		-3.0		-4.8	AVG. LOSS		-2.1		1.2		-2.7		-4.8

Table 3. Worst subtype pairings in terms of loss in each of the four performance metrics for the reference subtype for LHP (left) and RHP (right).

action). However, from Table 3, we find that KC pairs poorly with SI (dissimilar speed and dissimilar movement in both directions) and SL (similar speed and most similar movement other than CU) for RHP. Unlike for RHP, the KC and SL2 or KC and SL3 is a positive pairing (very different vertical movement) for swinging strike rate for LHP. For LHP, KC and FT has a negative pairing (very different horizontal and vertical movement and very different velocity) and plays down for swinging strike rate and ground ball rate.

For RHP, slider subtypes have various KC subtypes that pair well for increasing swinging strike rate and increasing ground and fly balls, whereas SI2 and SI3 reduce exit velocity. For LHP, sliders when paired with cutters, especially FC2 or FC3, seemed to reduce swinging strike rate and induce harder contact.

To accentuate the value of this pitch pairing research there are a couple of examples of one pitch in isolation having different performance than when paired with others. One example of a mediocre pitch having profound pairing impact is RHP SI2, which has only a 5.5 percent swinging strike rate and a decent ground ball rate (38.2). However, there are 17 instances where SI2 helps another subtype to have improved performance. For example, SI2 greatly helps slider with reducing hard contact. Conversely, RHP FS1 with low speed and low spin is somewhat of a black hole of pitch pairing, having 16 instances where it worsens the performance of a subtype. RHP, slider subtypes have various KC subtypes that pair well for increasing swinging strike rate and increasing ground and fly balls, whereas SI2 and SI3 reduce exit velocity. For LHP, sliders when paired with cutters, especially FC2 or FC3, seemed to reduce swinging strike rate and induce harder contact.

EXTREMES IN PAIRING FREQUENCY AND PITCHER PERFORMANCE LEVEL

In this section, we exemplify these pairings with five examples: Aaron Nola, Corey Kluber, Jeff Samardzija, Aaron Bummer, and Sean Newcomb.

Aaron Nola throws two types of curve balls (CU1 and CU3). Both of these types of curveballs have the greatest increase in swinging strike percentage when they are paired with the type of sinker that Nola throws, SI3. In other words, the optimal pairing for CU1 and CU3 are both SI3 for RHP, with the latter curve ball having the greatest increase. When we compare the physical characteristics, CU3 is faster with greater horizontal break. Therefore, Nola could potentially work with a pitching coach to increase the horizontal move-

ment on his curve even further to generate even more swinging strikes.

Corey Kluber is a particularly good example of a positive pairing with CU2 and SI2. Despite below-average velocity on the sinker, Kluber has had great success with those two pitches. In particular, since Kluber's sinker is of the SI2 variety, it helped his curve ball play up from a pitch subtype average swinging strike rate of 13.9% to a remarkable 26.2% swinging strike rate.

Jeff Samardzija also throws multiple breaking pitches, but many of them have bad outcomes. For example, CU2 and KC4 do not pair well for balls in play, KC4 and SL3 do not pair well for swinging strikes, and CU1 and FS1 do not pair well for whiffs or grounders. However, there is a silver lining here. In the last three springs, Samardzija has been working on a new breaking pitch, SL3, which has an excellent pairing with KC1 for RHP. He could work on his knuckle curve having more depth, which is the distinguishing feature of the KC1 and could take his SL3 to the next level, adding 2.9% of swinging strikes and 3.9% of pop-up rate.

Aaron Bummer and his short sample of one season with above-average results might well represent how minor league or amateur pitch tracking systems might pay great dividends. A good pitch pairing is the SI3 with the FF1 for LHP, which allows greater swinging strikes. He throws the four-seamer high in the zone following his sinker to get batters to swing under it. The pairing certainly contributed to the White Sox extending him for five years and \$16M even with his limited MLB success. With FF1 in his arsenal, his SI3 swinging strike rate was 9.0%, 1.8% more than the average SI3.

Sean Newcomb is another left-handed pitcher with a fairly short track record but has some interesting pairings that could produce high levels of performance in the future. Specifically, Newcomb pairs a CU1 and FF3 (a pairing that is not shown in Table 2, but very positive nonetheless) that generates a very high rate of ground balls. Since he had FF3 in his arsenal, his CU1 had a 56.0% ground ball rate, 13.2% above the average CU1. Despite his iffy command, it looks like Newcomb may return to the rotation this year, and that fly ball suppression may be part of the reason why.

CONCLUSION

In this paper, we used k-means clustering and the elbow method to classify pitch subtypes from previously-labeled MLB Statcast pitch types. In doing so, we

understood the degree to which subtypes differ across a type and evaluated the effectiveness of pairing subtypes. Between the best and worst pairing of subtypes, we found that there is an average change of 3.5 percent swinging strike rate, a 2.4 MPH exit velocity, 3.3 percent pop-up rate, and 4.5 percent ground ball rate. Lastly, based on frequency of the best and worst pairings, we showed examples of pitchers and discussed their level of performance. We hope that this work leads to intuition on where to focus efforts with pitcher scouting, pitch design, and development. ■

Notes

1. Max Marchi, "Pitch classification revisited." <https://tft.fangraphs.com/pitch-classification-revisited>, July 2010, and, Ethan Moore, "An unsupervised approach for pitch classification," in *Proc. of Sabermetrics, Scouting, and the Science of Baseball (Saberseminar)*, 2019.
2. Harry Pavlidis, "What makes a changeup good?" <https://www.baseball-prospectus.com/news/article/21675/what-makes-a-good-changeup-an-investigation-part-three>, Aug. 2013.
3. Glenn Healey and Shiyuan Zhao, "Finding the optimal pitch distribution," in *Proc. of Sabermetrics, Scouting, and the Science of Baseball (Saberseminar)*, 2019.
4. Carson Sievert and Brian M. Mills, "Using publicly available baseball data to measure and evaluate pitching performance," 39–66, New York, New York, USA: Chapman and Hall/CRC, 2017.
5. Tirthajyoti Sarkar, "Clustering metrics better than the elbow-method," <https://towardsdatascience.com/clustering-metrics-better-than-the-elbow-method-6926e1f723a6>, Sep. 2019.
6. Jonah Pemstein, "A long-needed update on reliability," <https://blogs.fangraphs.com/a-long-needed-update-on-reliability>, Sep. 2016.