4-2020

Investigating Major League Baseball Pitchers and Quality of Contact through Cluster Analysis

Charlie Marcou
Grand Valley State University

Investigating Major League Baseball Pitchers and Quality of Contact through Cluster Analysis

Charlie Marcou

Introduction

The rise of sabermetrics, the quantitative analysis of baseball, has changed how baseball front offices operate, how prospects are evaluated and developed, and how baseball is played on the field. Stolen bases are on the decline, while strikeouts, walks, and homeruns have steadily increased. Hitters care more and more about their launch angle and pitchers have started using high speed cameras to analyze their movement. Despite these changes, there are still many areas that need investigation. This paper seeks to investigate the quality of contact that a pitcher allows. Not much is currently known about quality of contact, but if factors determining quality of contact could be determined it could assist teams in identifying and developing pitching talent.

It's been argued that pitchers can't control the outcome of batted ball events (McCraken 2001), but more recent research suggests that pitchers can at least control whether batted balls are groundballs or flyballs as well as how hard the ball is hit (Tippet 2004, Fast 2011, Arthur 2015). A high number of ground balls might be an indicator of weak contact, but it comes at a cost of more baserunners (Zimmerman 2016). Spin rate is another potential indicator because low spin fastballs lead to less swing and misses and more ground balls, while the reverse is true for high spin fastballs (Day 2013, O'Connell & Marsh 2016). However, spin on individual pitch types have been found to not correlate with a pitcher's groundball rate (Long 2016). There is little difference in run-prevention when pitchers are at the extreme ends of the groundball-flyball spectrum (Clemons 2019). However, other research has indicated that a pitcher needs to have an extreme groundball rate to receive any special benefit from being groundball focused (Sarris 2019, Zimmerman 2015).

One problem with researching quality of contact is that pitchers are constantly making adjustments to their pitches and because of this hard hit rates don't stabilize (Zimmerman 2016).

A pitcher’s exit velocity is more consistent with a low number of balls in play than it is with larger samples (Carleton 2016). Essentially, the variables regarding quality of contact will vary over time due to a pitcher’s adjustments. Although lower velocity has been found to lead to worse results in quality of contact statistics (Zimmerman 2016), the constant adjustments makes it difficult to examine quality of contact for a long period of time.

The other problem is that the average quality of contact amongst pitchers has much less variation than batters. Pitchers do have control, but the batter is more important in determining the result (Fast 2011). The batters could have anywhere between two-thirds of the control (Fast 2011) to five-sixths (Arthur 2015).

Despite these problems that come with investigating what type of control pitchers have over contact allowed, one area to investigate is whether quality of contact is a repeatable skill. Furthermore, if it is a repeatable skill, then it is important to investigate what kind of benefit controlling contact allowed brings a pitcher. Along with this, groundball and flyball tendencies, and the types of pitches a pitcher throws will also be investigated.

Methodology

To investigate quality of contact I utilized data from MLB’s tracking system, Statcast. Statcast has only been around since 2015 so past research on quality of contact have only been able to use a few years of data, but it allows the tracking of previously unmeasurable statistics. Statcast accomplishes this by using high-resolution cameras and radar equipment in every MLB stadium; these cameras track the movements of the ball and players (Casella 2015). What Statcast measures is wide-ranging. This includes pitching, hitting, baserunning and defensive metrics.

I pulled all of the data from Statcasts' custom leaderboards page. The custom leaderboards organizes the relevant data by pitcher seasons and it can be customized to include any variable of interest. Due to the nature of the large amount of data and the website, I had to download data year by year and section by section. This was necessary because when selecting too much data at a time, the site would give an error that no data could be found. I also only downloaded the data of pitchers who had 50 plate appearances at minimum. I first downloaded data about exit velocity and quality of contact, then I went on to more general and pitch location statistics and finally variables about individual pitch types. These variables about a pitcher's arsenal were only available for the 2017-2019 seasons.

I merged each dataset together in R by using player names, years, player ages, and games played. I discovered that there were some observations in the newly created dataset that were exact copies of others and had them removed. I also had to reconvert some variables into numeric variables after the merges. After this, I created ERA using earned runs and innings pitched. Innings pitched was initially formatted so that .1 indicated a third of an inning and needed to be re-coded in order to accurately calculate ERA. This gave me 68 variables with 3,016 observations. Each observation is one pitcher season.

Table 1. Key Variables

Earned Runs	A run that a pitcher gave up that was not a result of an error made by a defensive player
Innings Pitched	The number of innings a pitcher has completed. This is measured by the number of outs gotten while a pitcher is on the mound.
Earned Run Average (ERA)	This is an average of earned runs per nine innings pitched.
Weighted On-Base Average (wOBA)	A statistic that attempts to weight the values of each outcome differently instead of treating all times on base equally. For example, homeruns and doubles are weighted more heavily than singles, and hits are weighted more heavily than walks. It is also readjusted to be on the same scale as On-Base Percentage (OBP).
Strikeout Rate (K%)	Measures how often the pitcher strikes out a hitter per plate appearance.
Walk Rate (BB%)	Measures how often the pitcher walks a hitter per plate appearance.
Average Exit Velocity	The average speed the ball is travelling after it hits and leaves the hitter’s bat.
Hard Hit Rate (Hard Hit %)	The rate of balls that had an exit velocity of 95 miles per hour or more.
Slugging Percentage (SLG)	Measures how well opposing hitters performed. It’s calculated by dividing total bases given up by at bats. Total bases is the number of bases a player has gained via hits. Homeruns are worth 4 bases, while singles are worth one.
Isolated Power (ISO)	Measures a hitter’s power similarly to SLG but does not include singles.
Barrel Rate	Measures the percentage of balls hit with a combination of exit velocity and launch angle that leads to at least a .500 batting average and 1.500 SLG.
Ground Ball Percent (GB%)	Measures the rate of ground balls per ball in play.
Fly Ball Percent (FB%)	Measures the rate of fly balls per ball in play.
Pop-Up Percent	Measures the rate of pop-ups per ball in play.
Fastball Percent	The percentage of pitches thrown that were a type of fastball.
Off Speed Percent	The percentage of pitches thrown that were an offspeed pitch.
Breaking Percent	The percentage of pitches thrown that were breaking pitches.

4-Seam Percent	The percentage of pitches thrown that were 4-seam fastballs.
Slider Percent	The percentage of pitches thrown that were sliders.
Changeup Percent	The percentage of pitches thrown that were changeups.
Curveball Percent	The percentage of pitches thrown that were curveballs.
Sinker Percent	The percentage of pitches thrown that were sinkers.
Cutter Percent	The percentage of pitches thrown that were cutters.
Splitter Percent	The percentage of pitches thrown that were splitters.
Knuckle Percent	The percentage of pitches thrown that were knuckleballs.
Pitch Average Speed	Measures the average speed of a pitch type. An example would be the average speed of a pitcher’s cutter.
Pitch Average Break	Measures the average break of a pitch type in inches. An example would be the average break of a pitcher’s slider.

Results

During initial exploratory analysis, I examined the effects of exit velocity and hard hit rate on some key variables. This was accomplished by examining the distributions of exit velocity and hard hit rate within high, medium, and low categories of the variables. Observations would be grouped into the low category if they were within the first quartile and grouped into high if they were within the fourth quartile. Observations in the second and third quartiles were grouped into a medium category. Figures 1-4 show that generally, pitchers with higher exit velocities and hard hit percentages tend to have worse performances than pitchers with low exit velocities in terms of general run prevention and extra base hits. Figures 5 and 6 show that the distributions of exit velocity and hard hit rate are not different depending on ground ball rate. This seems to be at odds with some previous research (Zimmerman 2016), but the difference in findings may be due to not having enough categories of ground ball percentage.

Figure 1.

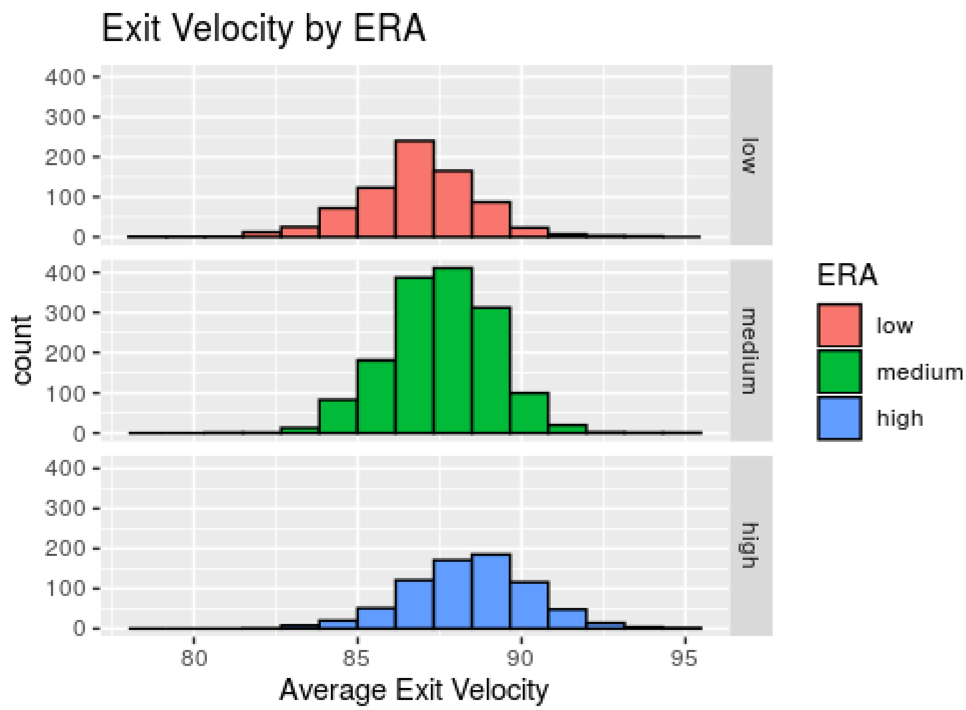


Figure 2.

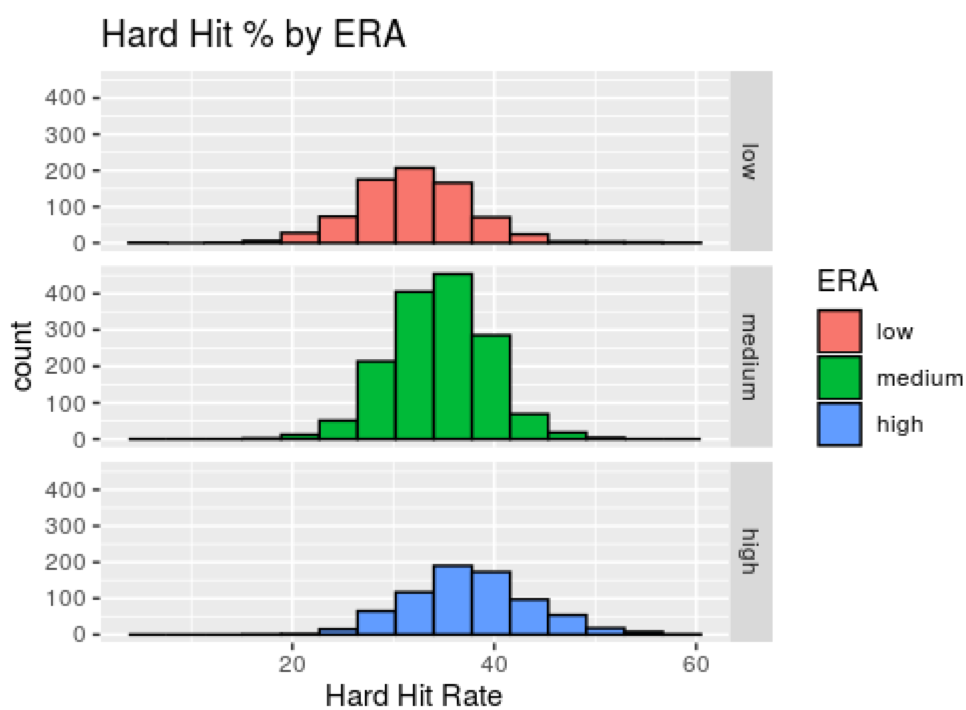


Figure 3.

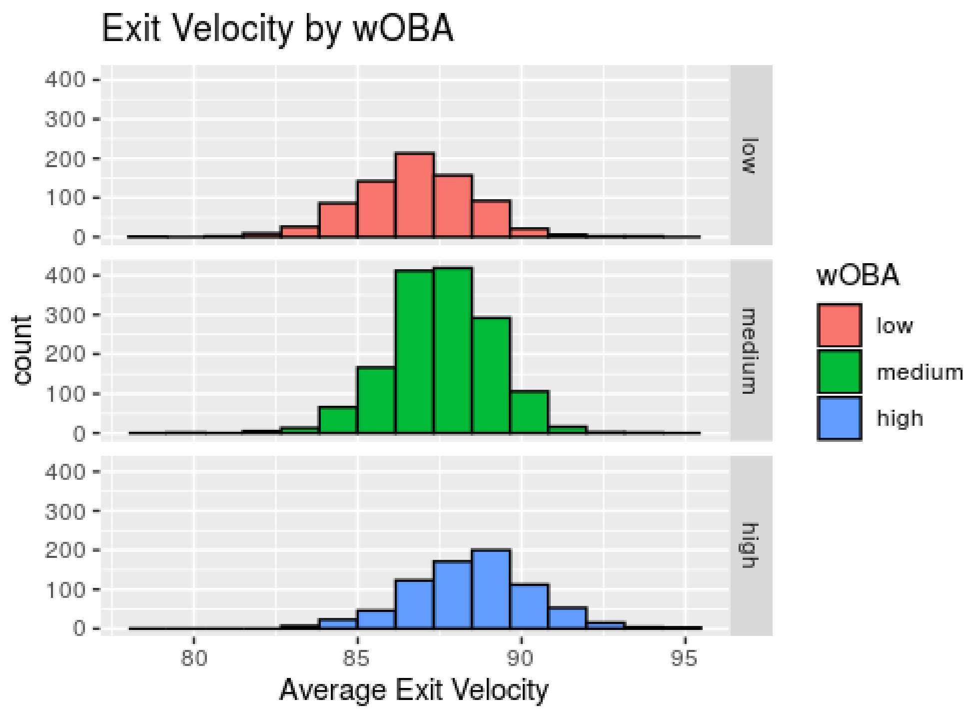


Figure 4.

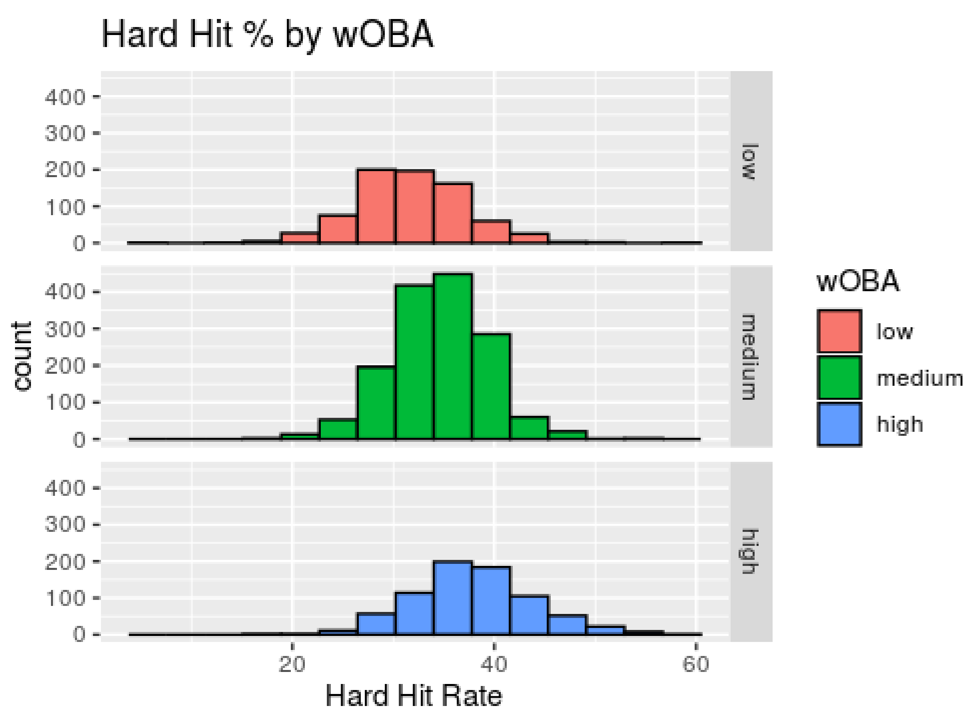


Figure 5.

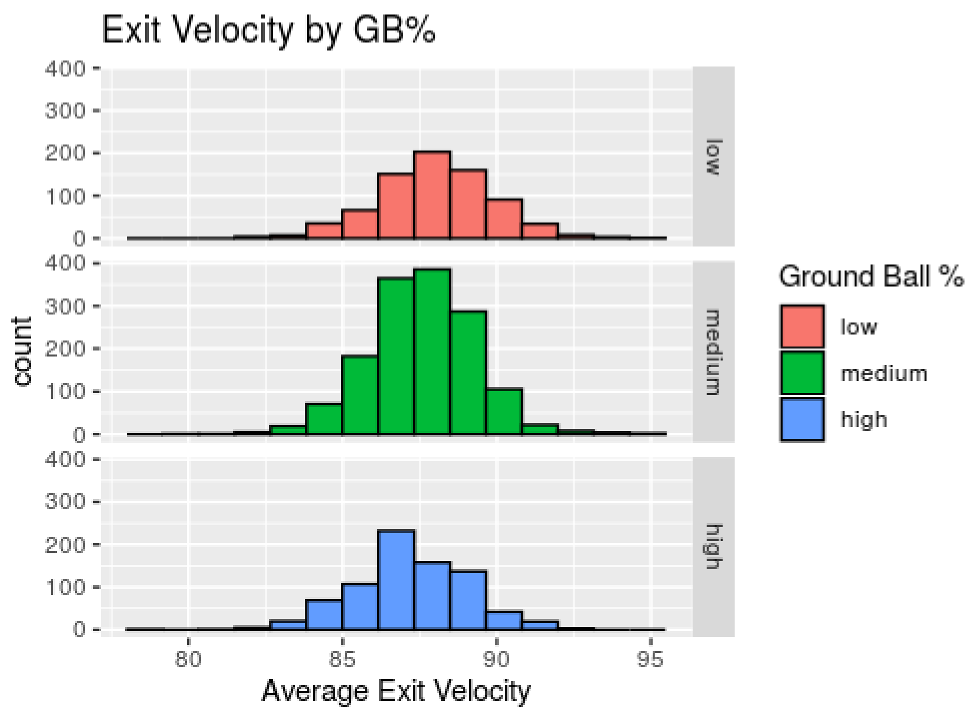
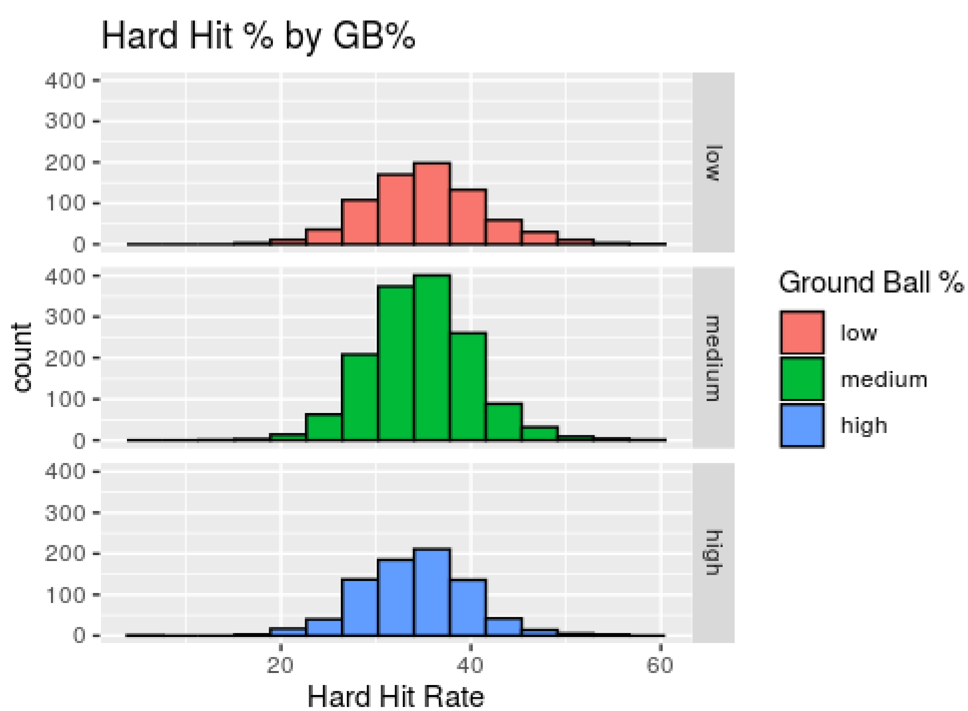


Figure 6.



In order to determine if a pitcher's arsenal explained anything about contact allowed, I decided to also examine the distributions of variables regarding pitch type. Fastballs, offspeed, and breaking pitches were examined as groups. They were grouped because not every player throws each pitch and by grouping there were fewer missing observations. However, grouping them like this might hide some effect a pitch type might have. For example, 4-seam fastballs and sinkers will move in different directions and might not have the same result.

In general, the distributions were similar for each category of exit velocity and hard hit rate. However, for fastball speed and the breaking pitch variables the distributions were slightly closer to being uniform in the low exit velocity group. The pitchers in the other two groups had closer to normal distributions. Figures 7-10 illustrate this. Table 2 shows the standard deviations of each variable by the exit velocity group of the pitcher. For example, the standard deviation of breaking pitch percentage for low exit velocity pitchers is 2.79 standard deviations larger than for high exit velocity pitchers.

Figure 7.

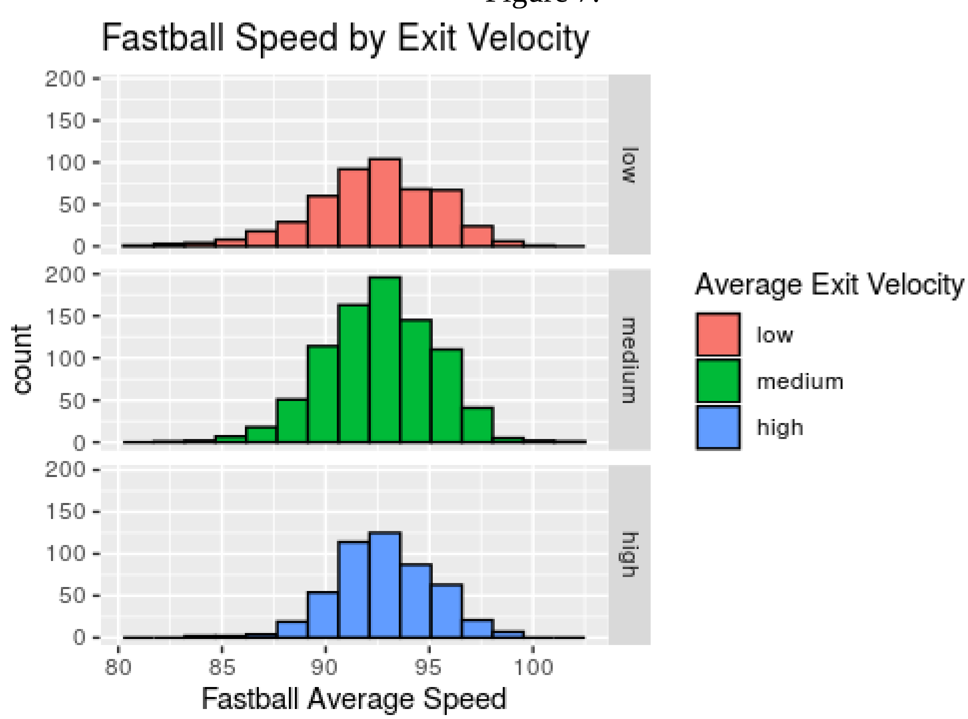


Figure 8.

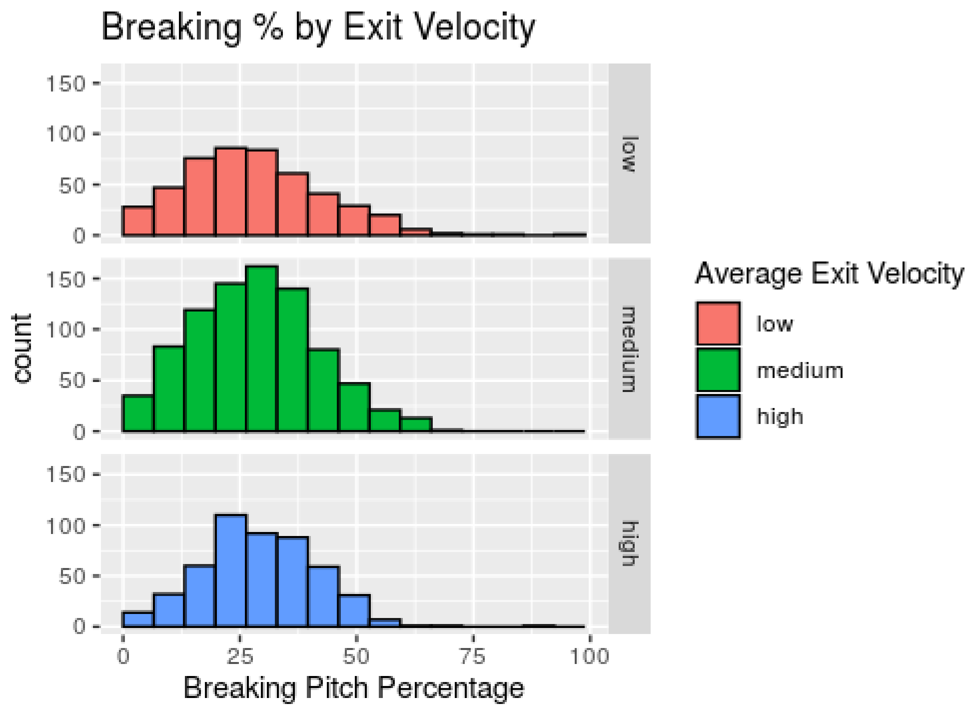


Figure 9.

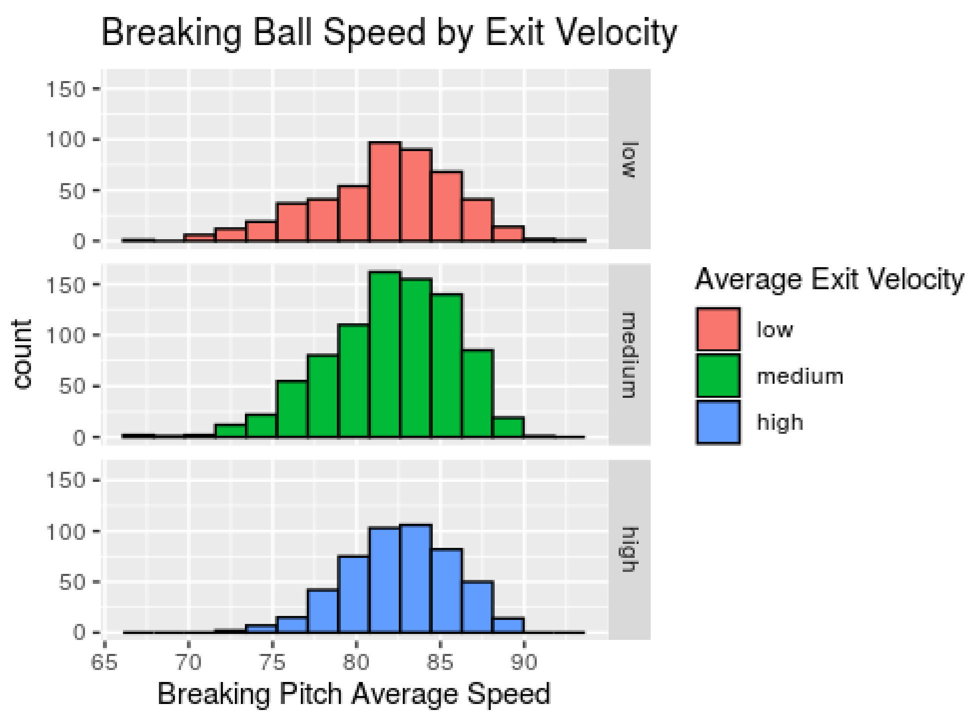


Figure 10.

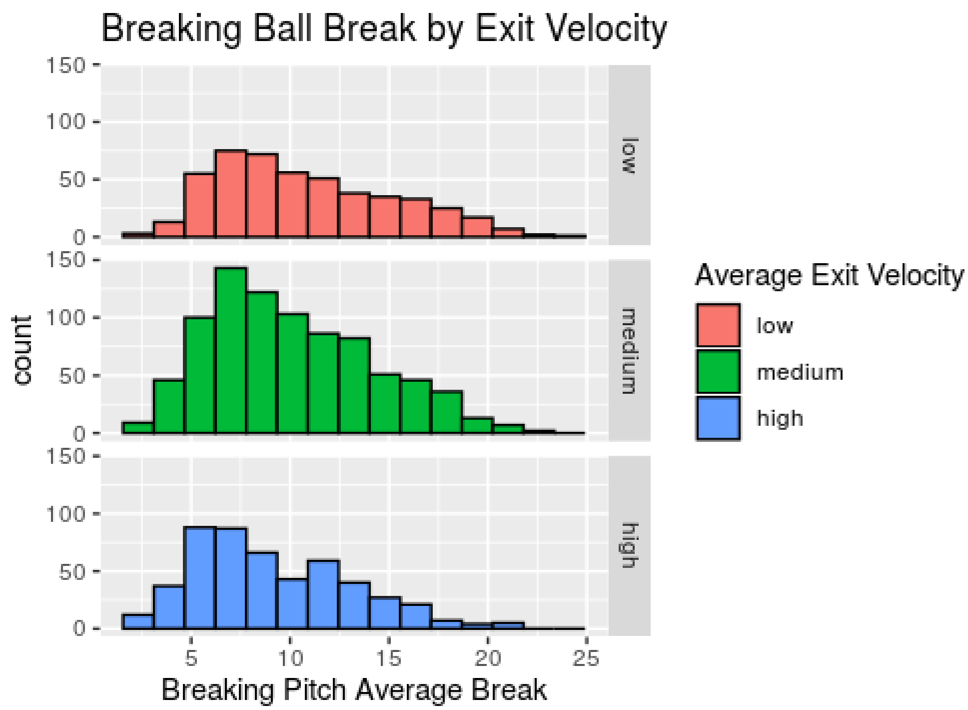


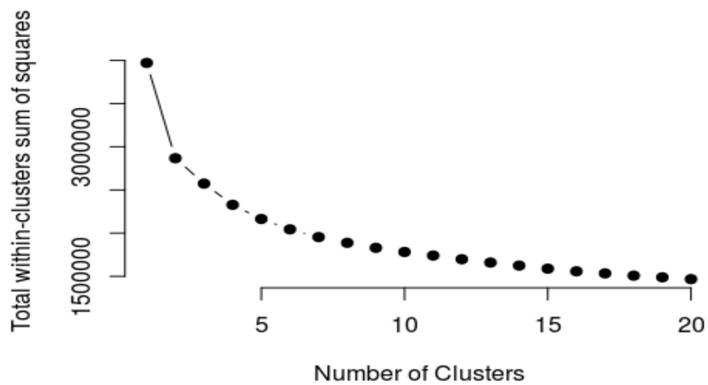
Table 2. Standard Deviations of Pitch Type Variables by Exit Velocity

Variable	Low Exit Velocity	Medium Exit Velocity	High Exit Velocity
Fastball Average Speed	3.07	2.61	2.40
Breaking Pitch %	15.05	13.33	12.26
Breaking Pitch Average Speed	4.21	3.73	3.17
Breaking Pitch Break	4.49	4.14	3.99

After establishing that there is an advantage to lower hard hit rates and lower exit velocities, I decided to move forward with a cluster analysis. The purpose of clustering is to group or cluster similar pitchers together. If pitchers in a group tend to have different exit velocities than another group, then clustering allows us to look at what is different about those pitchers. If a pitcher appears in a cluster that has low exit velocities multiple times, then this might be evidence that quality of contact is a repeatable skill. Additionally, since there are so many variables and pitcher seasons in the dataset, clustering helps make more sense of the data.

I clustered the data using all variables having to do with quality of contact, as well as variables that had to do with pitch location, strikeouts, and walks amongst others. I used K-means clustering, which uses an algorithm to sort data points into a number of specified groups. Each data point gets sorted into the cluster with the nearest mean values. Because K-means requires a predetermined number of clusters to use, I used the elbow method. The elbow plot in Figure 13 doesn't have an incredibly clear elbow, but marginal gains in explained variance seem to drop after 6 clusters. Although I didn't use it for any analysis, I did use hierarchical clustering to confirm that six clusters would be appropriate. The initial cluster assignments are random so I specified that 50 different starting clusters should be tried and that the one with the lowest variation would be used.

Figure 11.



Based on the cluster means shown in Table 3, these 6 clusters could be characterized as: high K and high BB pitchers, GB pitchers, FB with high K, FB with low K, other, and low K low BB pitchers. Table 4 shows 3 examples of pitchers that appear multiple times within each cluster.

A few conclusions can be drawn from the table of cluster means:

- Cluster 1 (high K, high BB) gets the best results on average with the lowest mean ERA, wOBA and SLG.
- Cluster 2 (GB) performs the best in terms of quality of contact with the lowest mean average exit velocity, hard hit percent, and barrel rate. This conflicts with what figures 4 and 5 suggested.
- Cluster 2 (GB) and Cluster 3 (FB, high K) are tied for the second best in ERA. Cluster 2 outperforms Cluster 3 in terms of quality of contact, but Cluster 3's high K% likely explains the good performance.
- Cluster 3 (FB, high K) and 4 (FB, low K) have similar hard hit percents. They also have the highest FB%, lowest GB% and highest barrel rates.

- Cluster 1 (high K, high BB) and Cluster 2 (GB) have markedly lower barrel rates, the highest GB% and are both in the top 3 for lowest FB%.
- Cluster 5 (other) is the largest cluster with each cluster mean being close to the grand mean for each variable.
- Cluster 6 (low K, high BB) has the worst mean ERA, wOBA, ISO and SLG. It doesn't have the worst performance in the quality of contact variables but has a markedly lower K% than all other groups.

Table 3. Cluster Means of Key Variables

	N	K%	BB%	GB%	FB%	Avg Exit Velocity	Barrel Rate	Hard Hit Percent	ERA	wOBA	ISO	SLG
Cluster 1	423	27.89	9.90	48.07	20.80	87.32	5.83	33.70	3.96	.296	.144	.371
Cluster 2	342	19.73	8.90	58.86	14.23	87.11	4.57	33.40	4.11	.311	.132	.387
Cluster 3	403	27.64	9.24	35.51	28.07	87.62	7.64	34.28	4.11	.306	.180	.410
Cluster 4	489	19.52	8.74	35.41	28.32	87.97	7.97	35.12	5.23	.341	.207	.471
Cluster 5	851	21.23	8.58	45.50	21.79	87.48	6.36	34.70	4.53	.324	.168	.428
Cluster 6	508	15.12	8.51	47.90	19.94	87.83	6.32	36.17	5.38	.351	.181	.469
	3016	21.55	8.91	44.81	22.37	87.57	6.51	34.67	4.60	.324	.171	.427

Table 4. Examples of Pitcher Seasons in each Cluster

Cluster	Pitcher	Seasons	Pitcher	Seasons	Pitcher	Seasons
Cluster 1	Clayton Kershaw	2015-2017,2019	Andrew Miller	2015-2018	Noah Syndergaard	2015-2018
Cluster 2	Zach Britton	2015-2019	Dallas Keuchel	2015-2019	Mark Melancon	2015-2019
Cluster 3	Pedro Baez	2015-2019	Max Scherzer	2015-2019	Justin Verlander	2016, 2018-2019
Cluster 4	Mike Fiers	2015, 2018-2019	Jake Odorizzi	2015-2018	Anibal Sanchez	2016-2017, 2019
Cluster 5	Zach Greinke	2015-2016, 2018-2019	Cole Hamels	2016-2019	Jon Lester	2015-2018
Cluster 6	Alex Cobb	2016-2019	Wade Miley	2015-2017	Jose Urena	2015-2019

While the cluster means of average exit velocity and hard hit rate are different, the distributions are very similar. Based on the distributions of the variables within each cluster, some of the important variables in the clustering are K%, GB%, FB%, and Pop-Up%. Pitchers in Clusters 1 and 3 tended to have higher strikeout rates, which is shown in Figure 14. Meanwhile, Figure 15 show that pitchers in Cluster 2 had higher ground ball rates than the other clusters. Figures 18-21 didn't show that there was a huge difference in ERA, wOBA, SLG, or ISO by cluster. But in general, Clusters 4, 5, and 6 performed slightly worse than the first 3 clusters. This might indicate that if a pitcher doesn't have a high K%, having a high GB% may be a way to make up for this. Additionally, Figures 16 and 17 show that clusters 3 and 4 have a similar FB% and Pop-Up%, but the high K% in Cluster 3 might explain the difference in results.

Figure 12.

Figure 13.

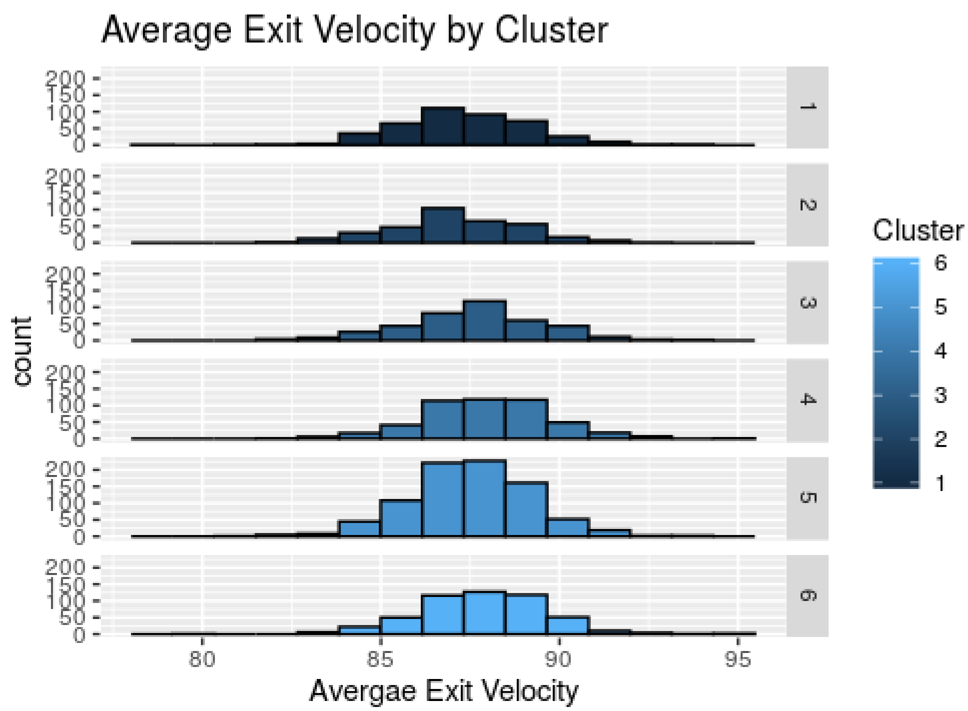


Figure 14.

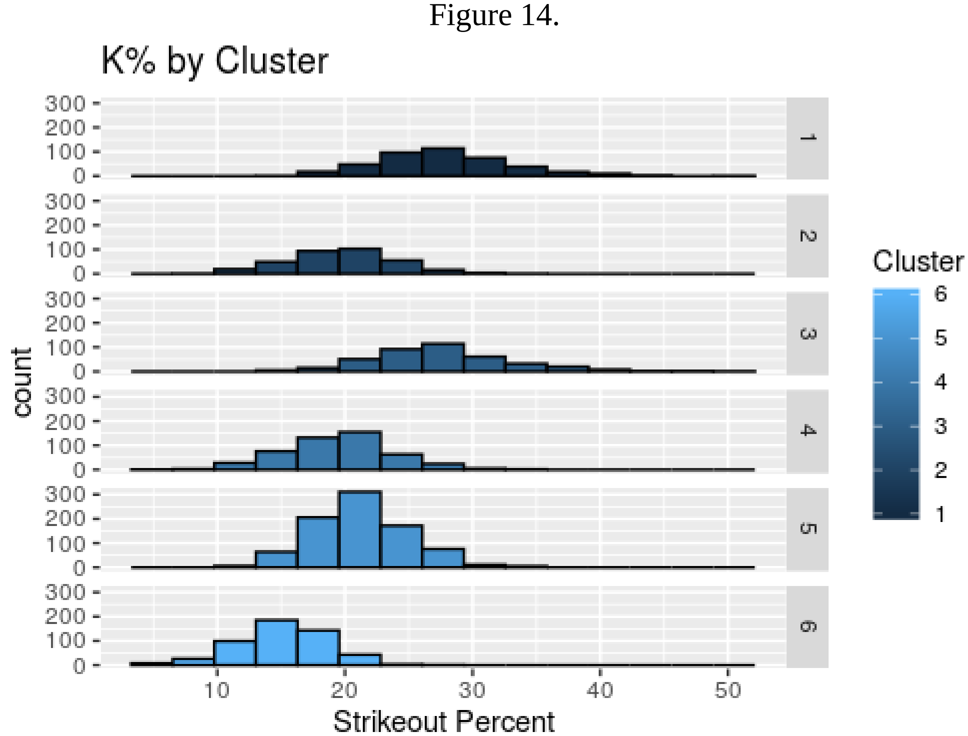


Figure 15.

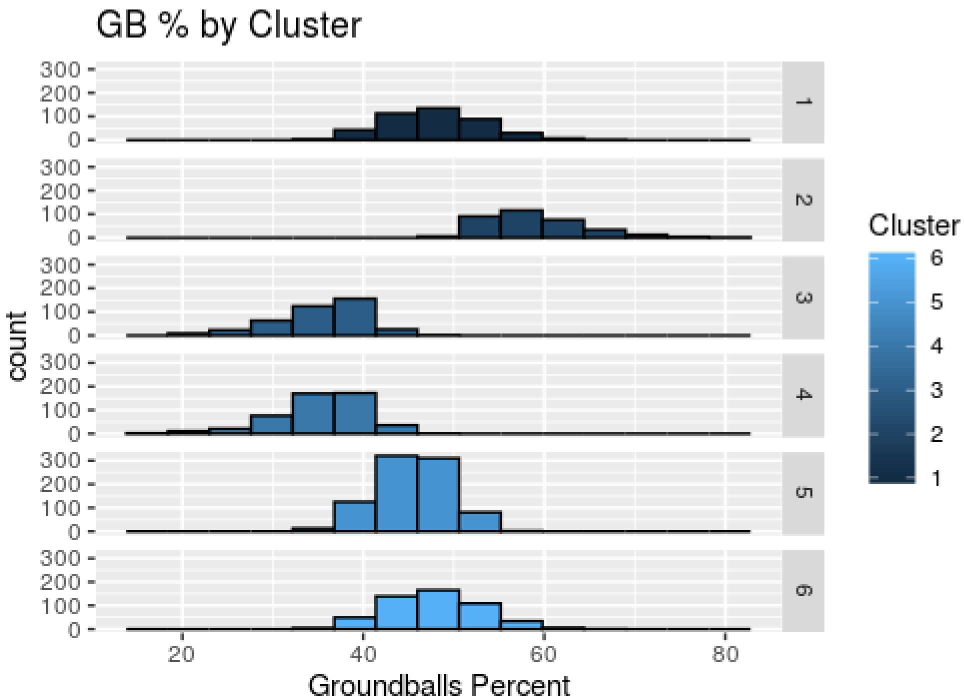


Figure 16.

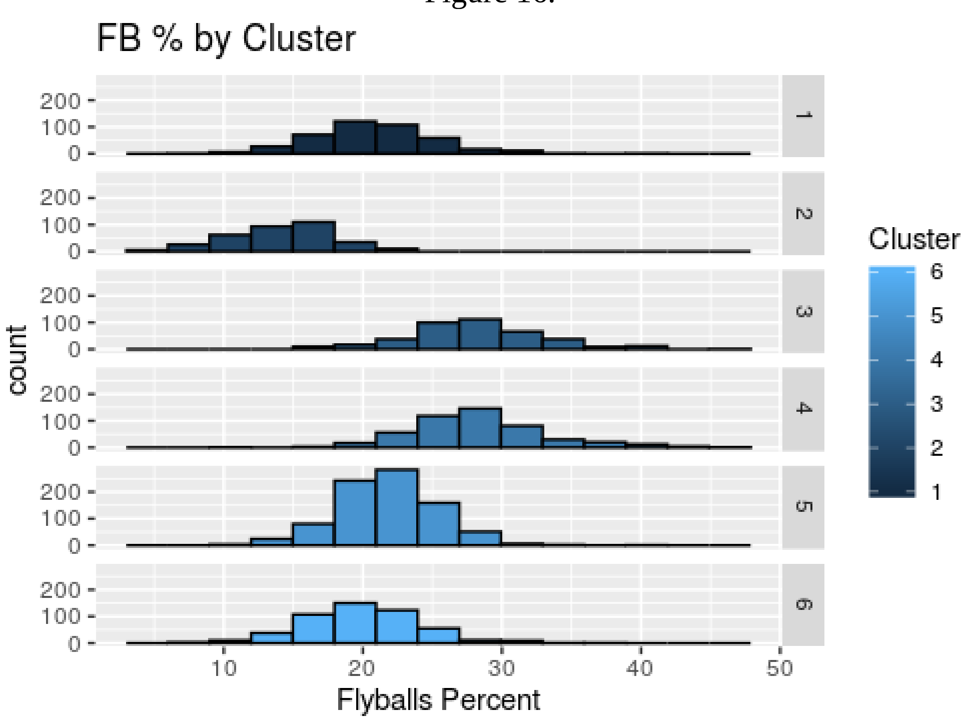


Figure 17.

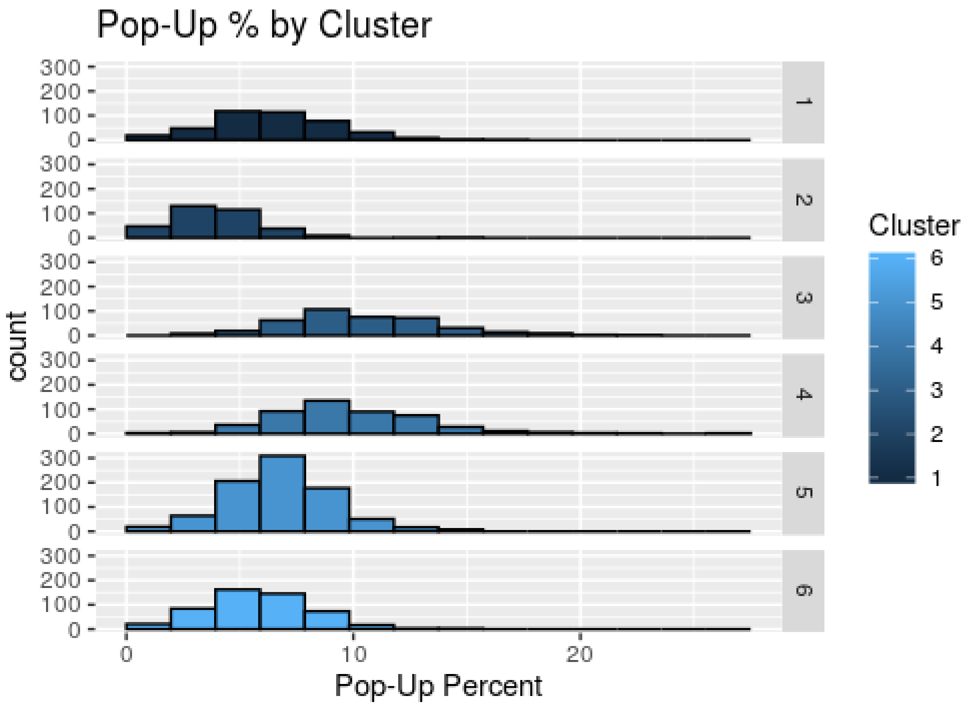


Figure 18.

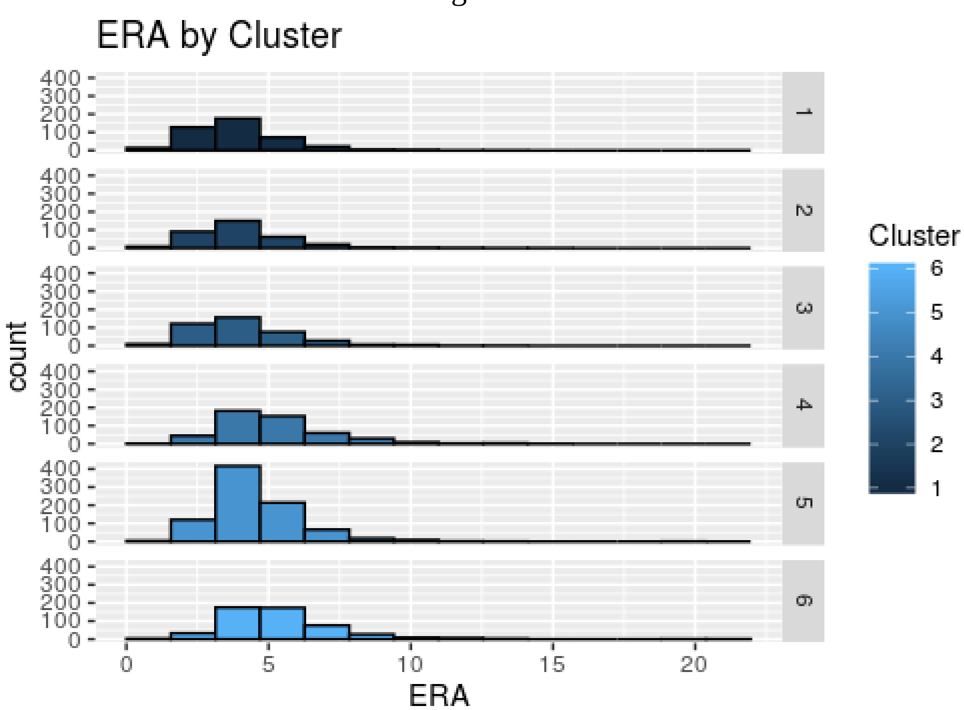


Figure 19.

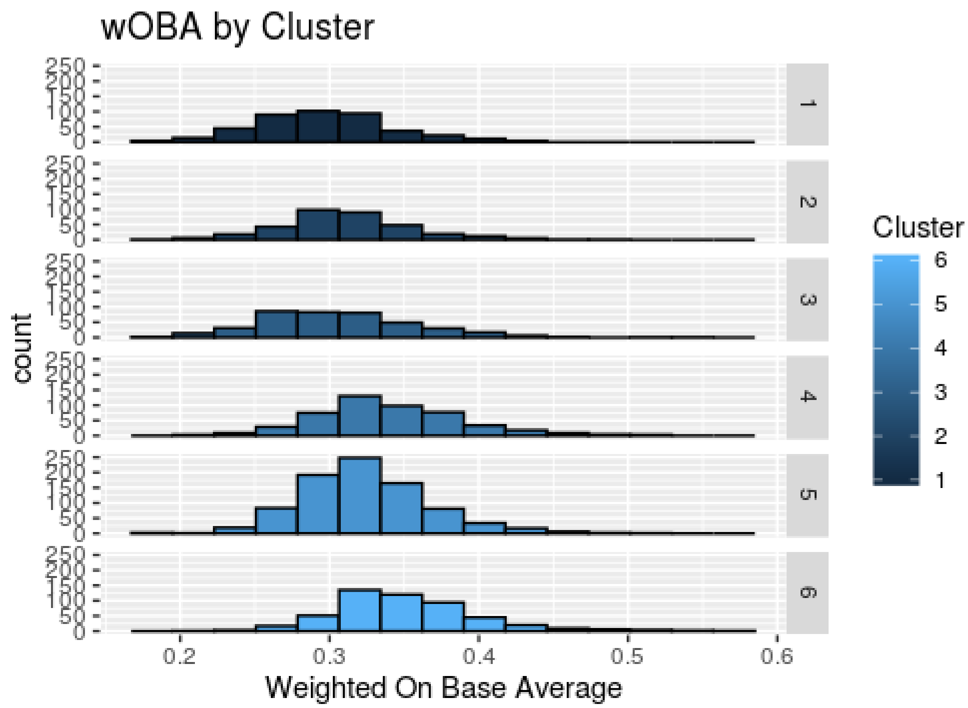


Figure 20.

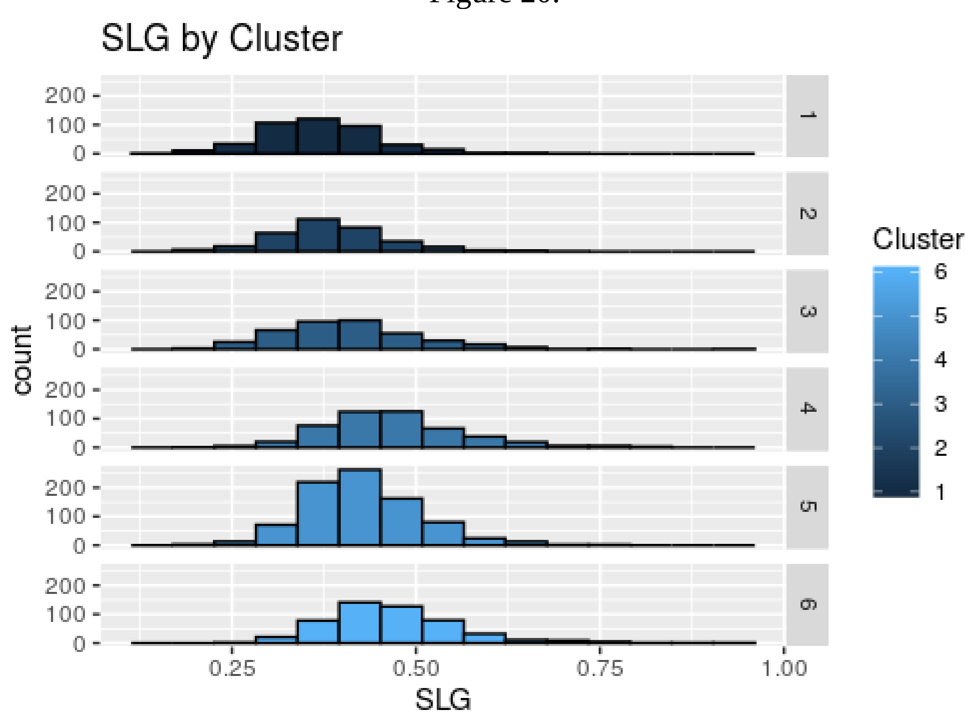
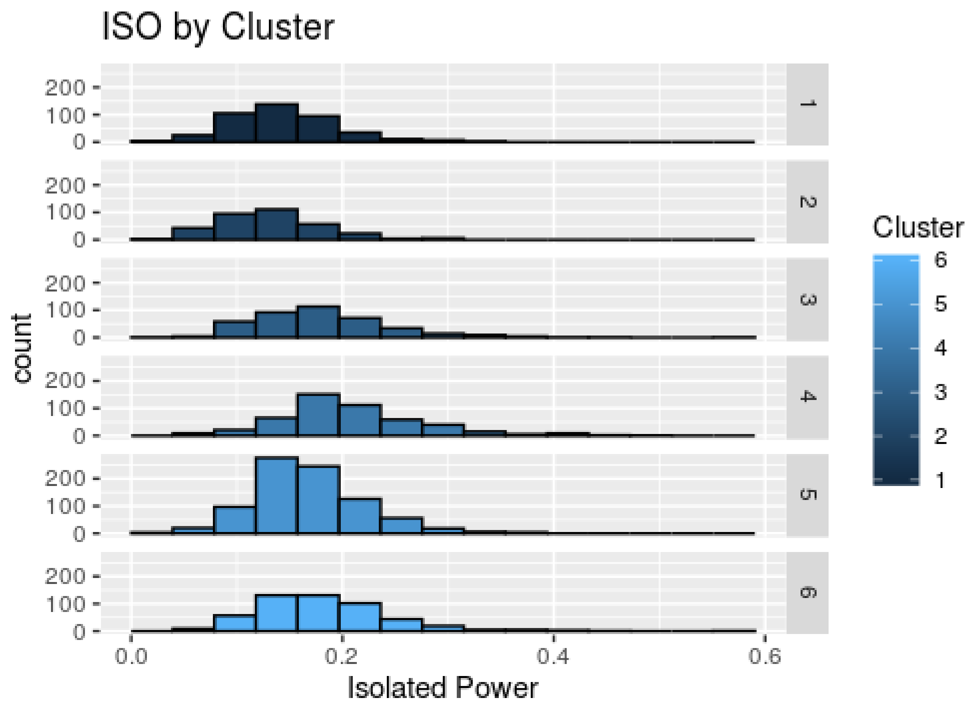


Figure 21.


ISO by Cluster
count
200
100
0
1
2
3
4
5
6
0.0
0.2
0.4
0.6
Isolated Power
Cluster
6
5
4
3
2
1

After this initial clustering I re-clustered, but this time instead of using the quality of contact variables, pitch arsenal variables were included. The other variables involved in the first clustering were still used as well. These pitch arsenal variables only included the percentage of each pitch type thrown. Velocity, spin and break were not included because not every pitcher threw each pitch. Since the pitch arsenal data was only available from 2017-2019, this clustering has a smaller sample.

Tables 5 and 6 show the 6 different cluster means. Table 5 shows that for the most part pitchers where separated by their pitch mix. The clusters can be categorized as: Cutterballers, Diverse Pitch Mix pitchers, Sinkerballers, 4-seam reliant pitchers, Fastball-Slider pitchers, and Sinker-Slider pitchers. Table 7 shows 3 examples of pitchers that appear multiple times within each cluster.

A few conclusions can be drawn from these cluster means:

- Cluster 1 (cutterballers) was the smallest cluster but had the best results with the lowest exit velocity, hard hit %, ERA, wOBA, and SLG.
- Cluster 6 (sinker-slider pitchers) had the next best mean results in ERA, wOBA, and SLG with the second highest GB%.
- Cluster 3 (sinkerballers) had the highest GB% and lowest barrel rate and ISO, but was about average in terms of exit velocity, hard hit rate and ERA.
- Cluster 4 (4-seam reliant pitchers) had the least diversity in their pitches and had an above average mean ERA, but Cluster 2 had the most diversity in their pitches and had an even higher ERA. This might suggest some specialization in pitch type is good.

- Cluster 4 (4-seam reliant ptichers) had the highest FB% and highest ERA, but Cluster 3 (sinkerballers) had the lowest FB% and an average ERA. This might suggest only extreme FB% that are on the high-end impact overall performance.

Table 5. Cluster Means of Key Variables when clustering by Pitch Arsenal

	N	K%	BB%	GB%	FB%	Avg Exit Velocity	Barrel Rate	Hard Hit Percent	ERA	wOBA	ISO	SLG
Cluster 1	156	22.25	8.44	44.92	22.59	87.02	6.49	33.62	4.23	.313	.165	.414
Cluster 2	352	19.27	8.72	45.25	22.63	87.30	6.68	34.76	4.89	.333	.180	.446
Cluster 3	262	18.33	8.64	51.77	19.07	87.49	6.22	35.68	4.72	.328	.162	.430
Cluster 4	502	23.02	9.23	39.51	26.47	87.98	7.59	36.18	4.82	.327	.189	.441
Cluster 5	370	25.62	9.79	41.26	24.88	87.65	7.25	35.30	4.67	.320	.179	.424
Cluster 6	197	22.73	9.32	47.61	21..8	87.29	6.37	34.32	4.54	.318	.167	.416
	1839	22.06	9.10	44.04	23.48	87.56	6.93	35.24	4.71	.325	.177	.432

Table 6. Second Clustering Cluster Means of Pitch Percentages

	4 Seem %	Slider %	Changeup %	Curveball %	Sinker %	Cutter %	Splitter %	Knuckle %
Cluster 1	24.37	4.45	7.45	12.95	12.25	37.57	.725	0
Cluster 2	33.20	12.76	13.40	13.89	21.99	2.48	1.32	.709
Cluster 3	6.36	11.48	12.76	9.37	54.30	4.81	.693	0
Cluster 4	58.05	9.82	11.16	13.29	1.85	3.25	2.36	0
Cluster 5	47.96	35.84	6.04	4.02	4.56	.511	.856	0
Cluster 6	16.44	36.68	5.22	3.15	35.99	.935	1.36	0

Table 7. Examples of Pitcher Seasons in each Cluster

Cluster	Pitcher	Seasons	Pitcher	Seasons	Pitcher	Seasons
Cluster 1	Madison Bumgarner	2017-2019	Corey Kluber	2017-2019	Jon Lester	2017-2019
Cluster 2	Luis Castillo	2017-2019	Sonny Gray	2017-2019	Aaron Nola	2017-2019
Cluster 3	Jake Arrieta	2017-2019	Kyle Hendricks	2017-2019	Dallas Keuchel	2017-2019
Cluster 4	Josh Hader	2017,2019	Rich Hill	2017-2019	Craig Kimbrel	2017-2019
Cluster 5	Carlos Carrasco	2017-2019	Jacob deGrom	2018-2019	Chris Sale	2017-2019
Cluster 6	Patrick Corbin	2017-2019	Carlos Martinez	2017-2019	Pedro Strop	2017-2019

Figures 22-34 show the distribution of variables within each cluster. Figures 22 and 23 indicate that the distribution of hard hit rate and average exit velocity is more uniform than within the other clusters. The other distributions seem skewed towards more extreme results with more pitchers with higher hard hit rates and exit velocities. Figure 24 shows that within Clusters 3 and 6, the clusters with sinker reliant pitchers, the pitchers tend to get more ground balls. Meanwhile, the other clusters have similar distributions of groundballs with Clusters 4 and 5 getting slightly less. Figures 25-28 show that Cluster 1 tends to perform slightly better than the other clusters in terms of ERA, wOBA, SLG and ISO. These distributions show that Cluster 1 has fewer extreme results. The distributions of wOBA, SLG and ISO are also more uniform for Cluster 1 than for the other clusters. Figures 29-34 confirm what the cluster means of pitch percentages suggest about the types of pitchers found in each cluster.

Figure 22.

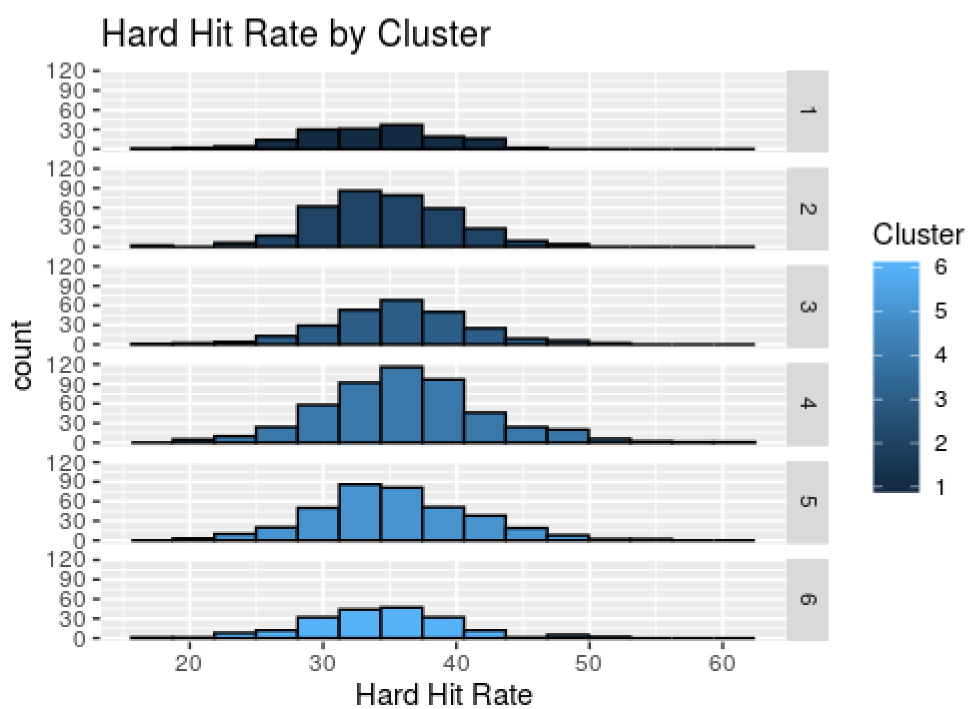


Figure 23.

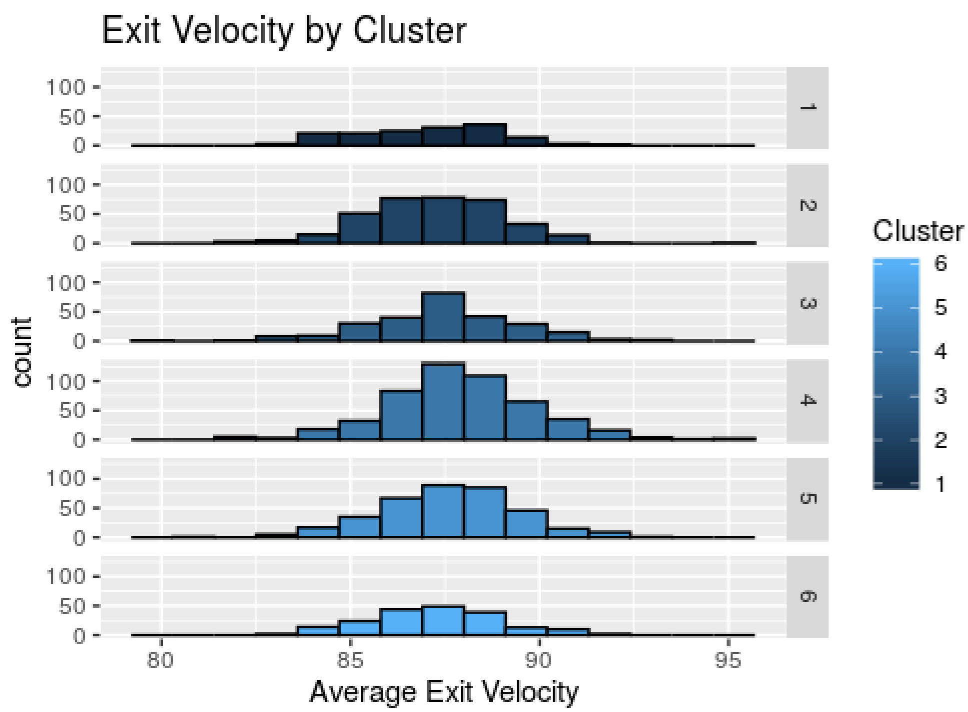


Figure 24.

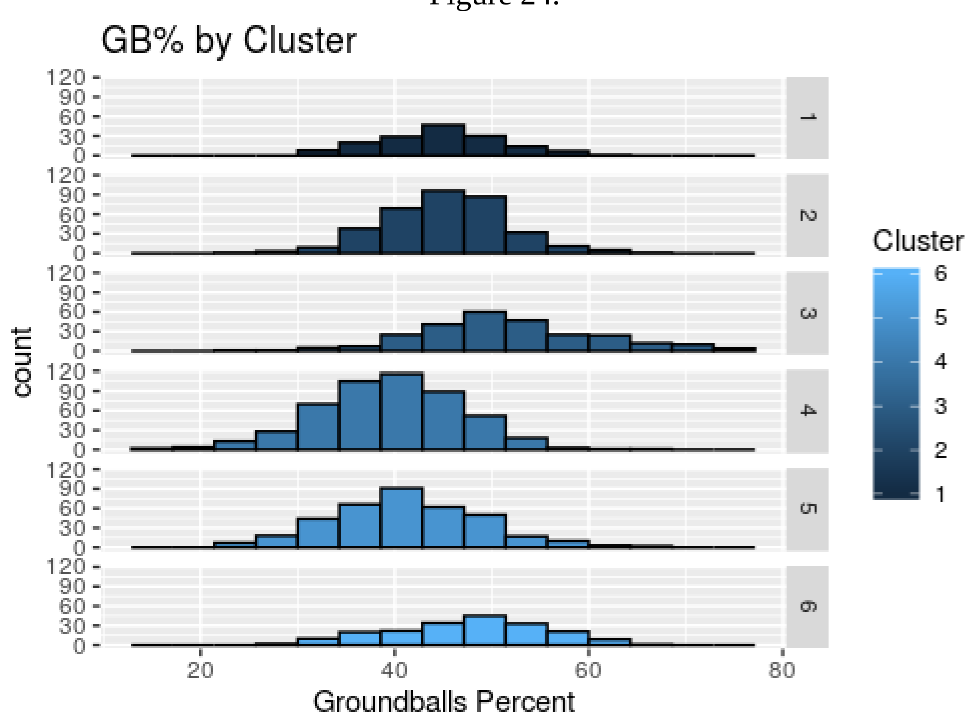


Figure 25.

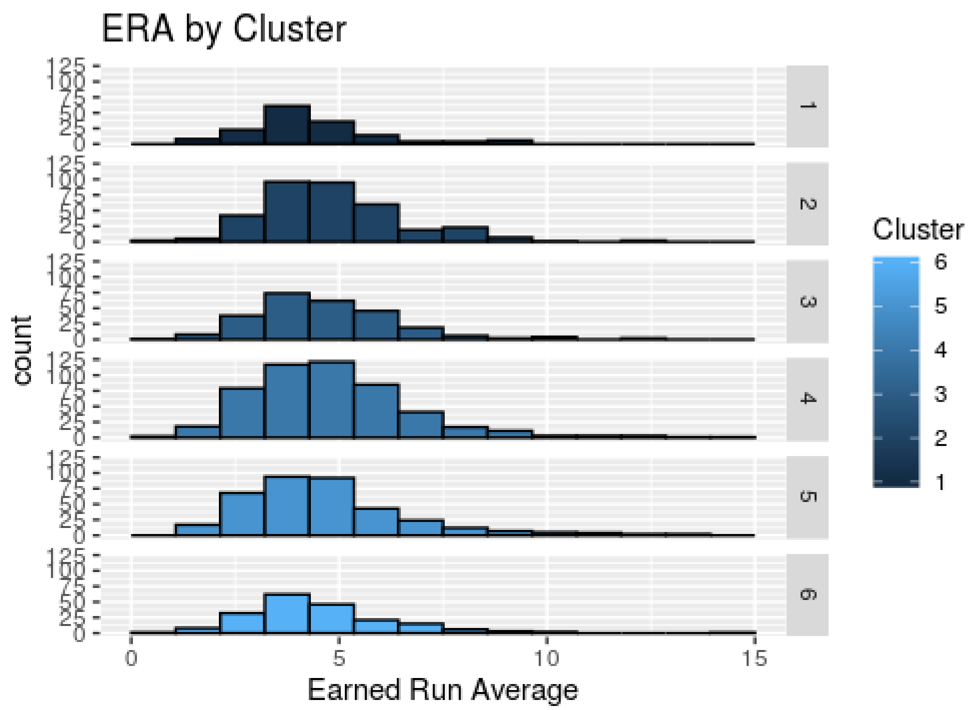


Figure 26.

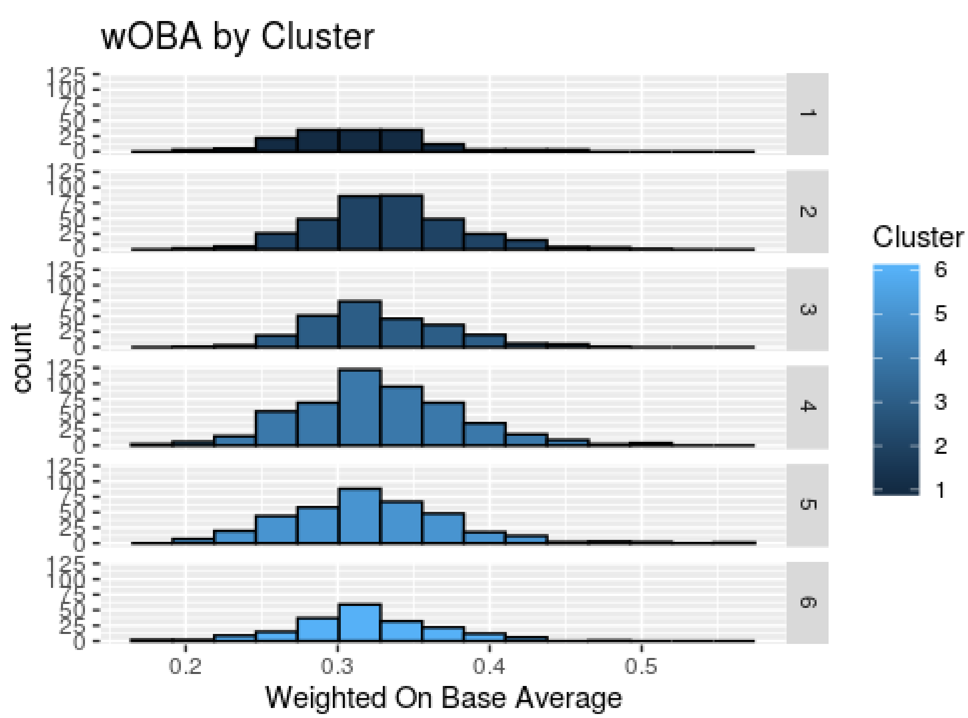


Figure 27.

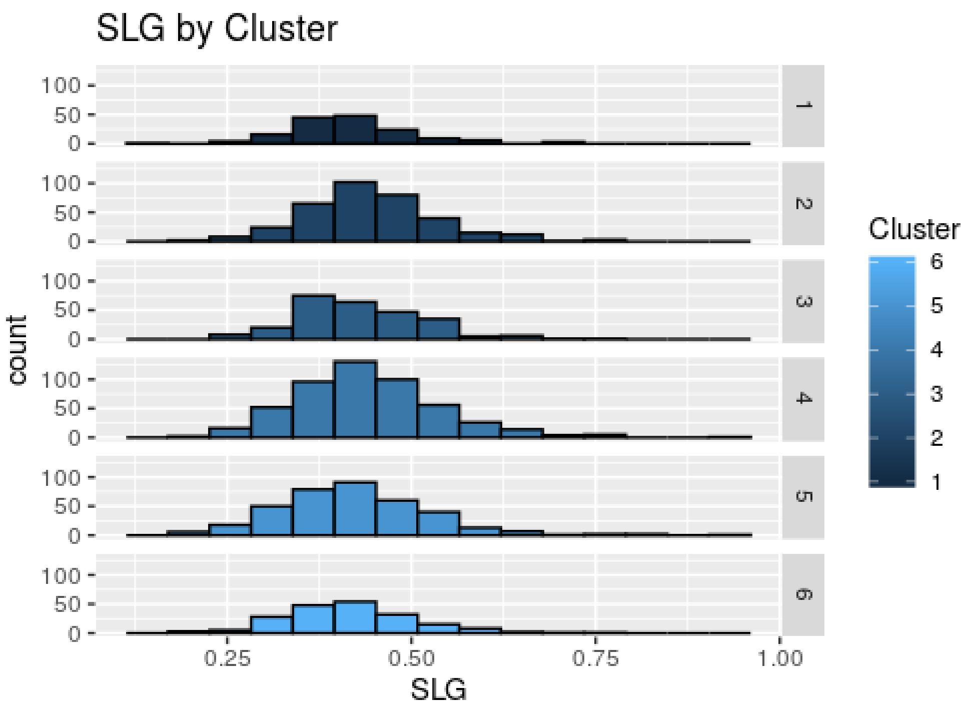


Figure 28.

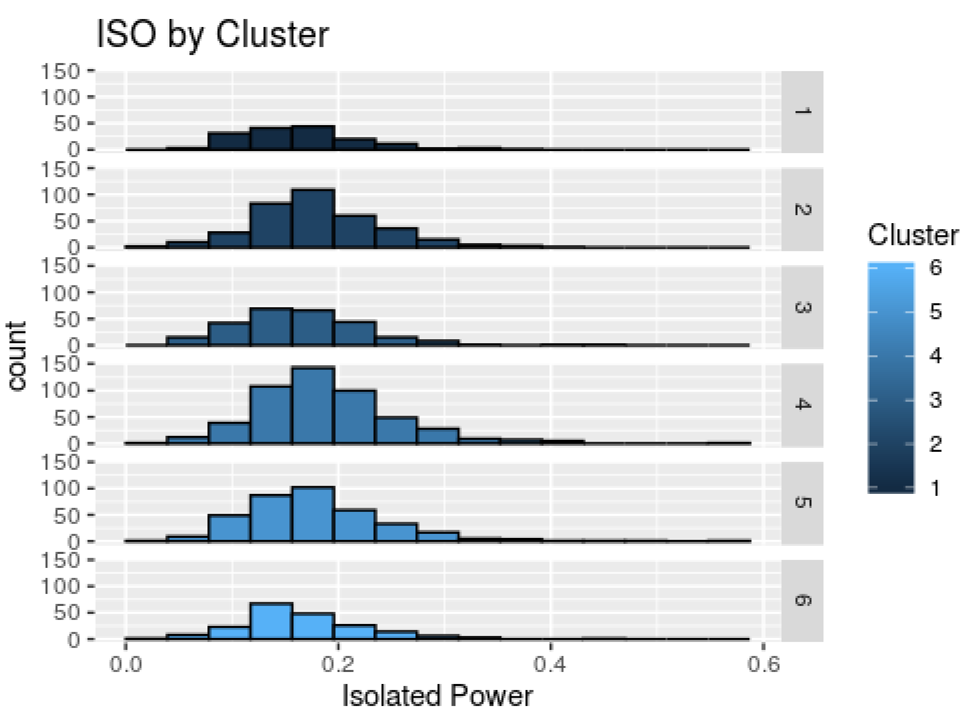


Figure 29.

Figure 30.

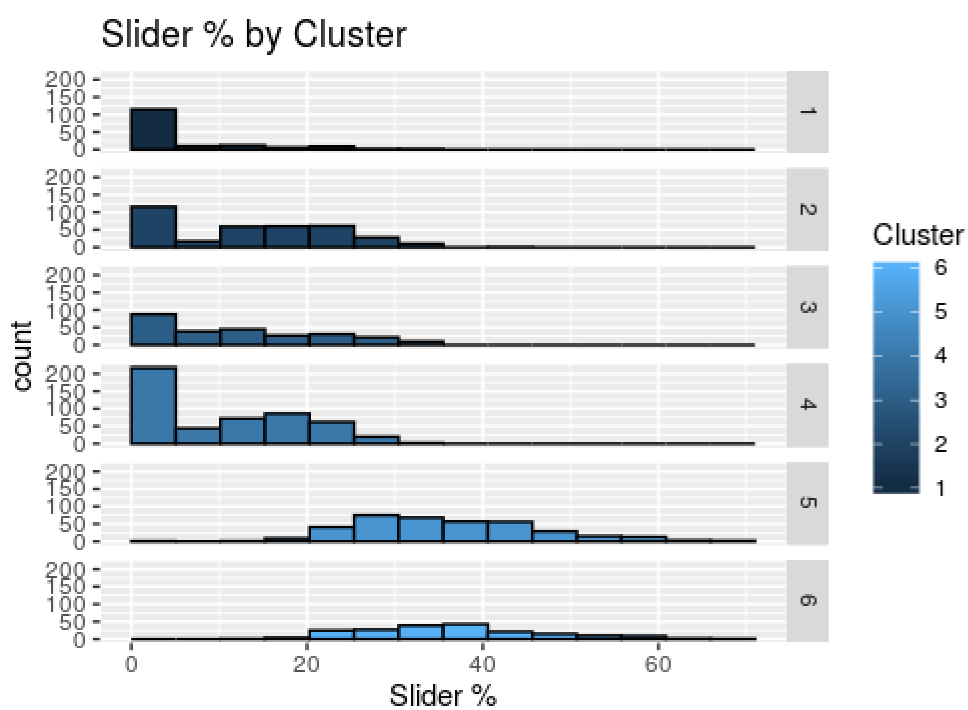


Figure 31.

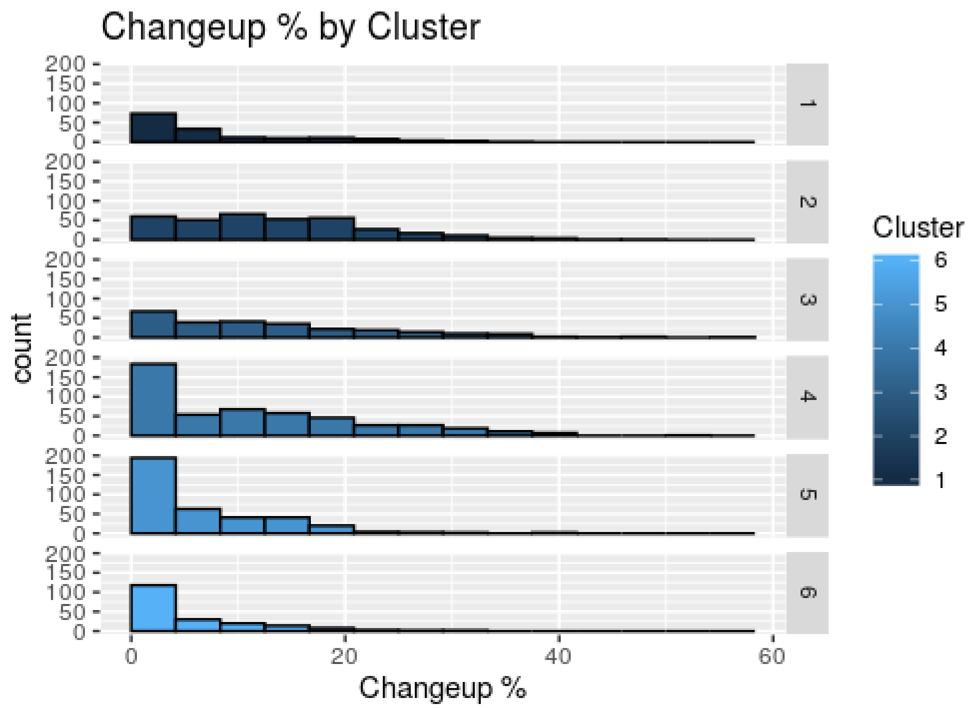


Figure 32.

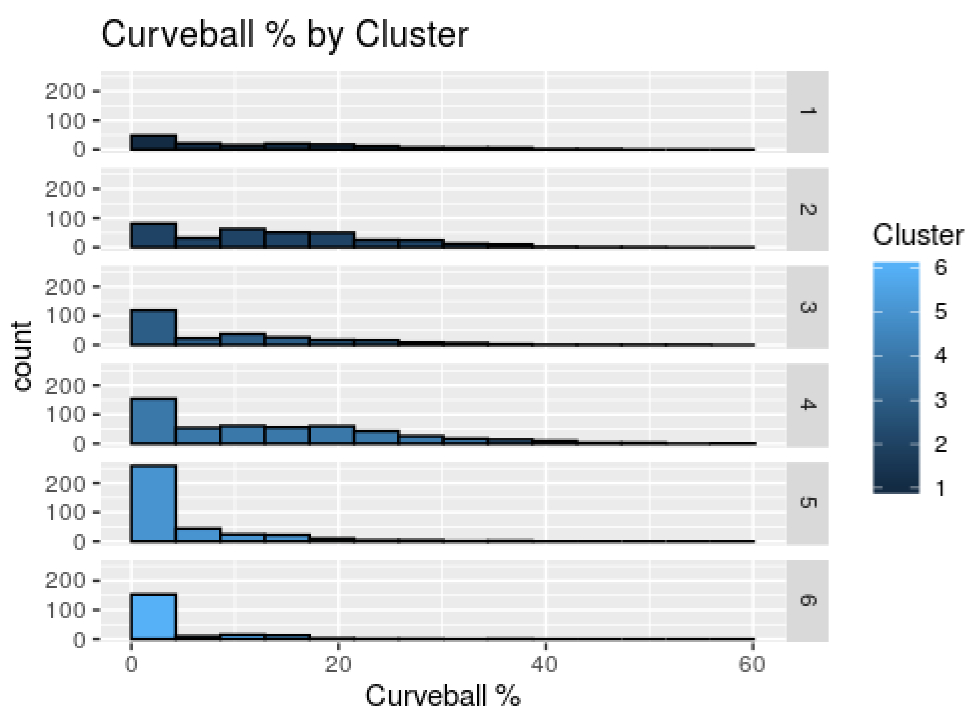


Figure 33.

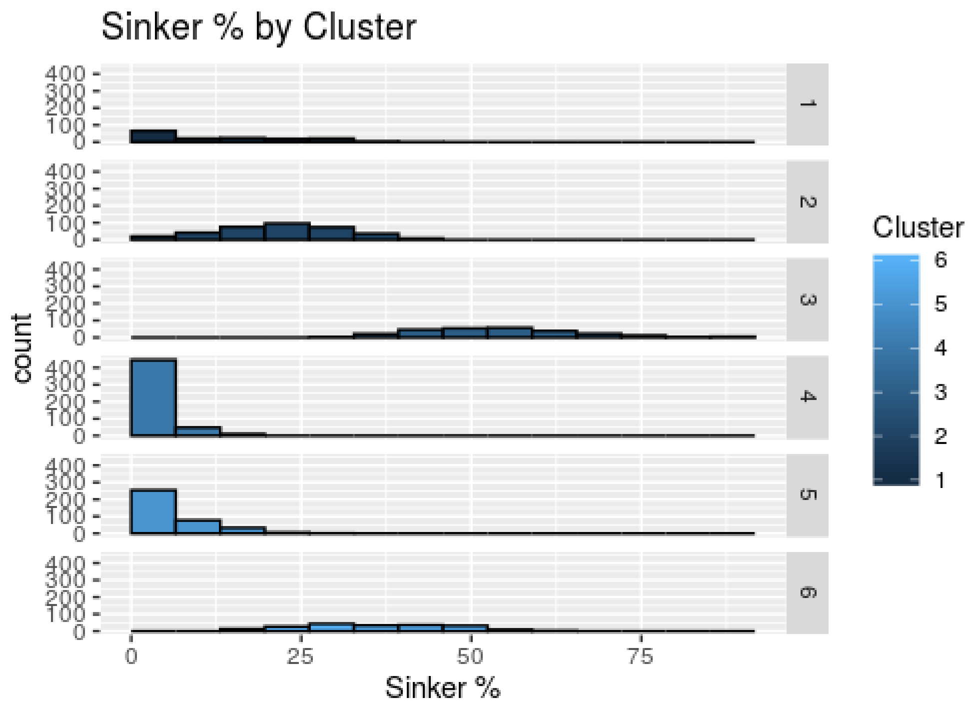
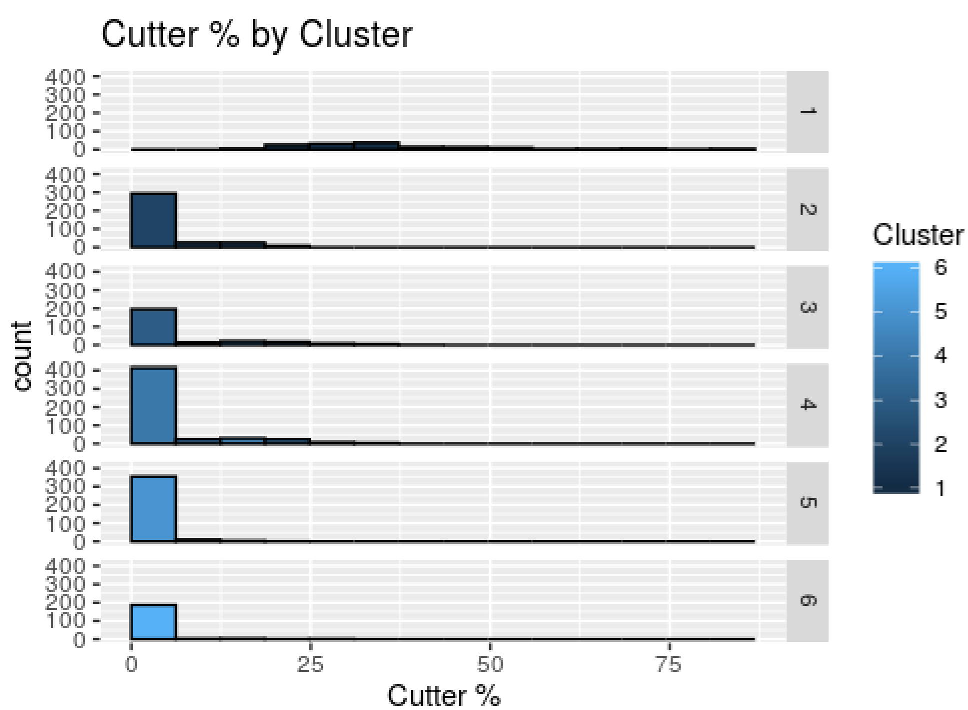


Figure 34.



Discussion

Due to the nature of clustering, no strong conclusions can be drawn from this analysis. Based off the first clustering, it seems as if there was little difference in the distributions of exit velocity and hard exit velocities in the clusters determined. There were differences in the cluster means and individual pitchers appeared multiple times within each cluster. But the lack of differences among the distributions does not support the idea that quality of contact is a repeatable skill.

However, this analysis does indicate that there is something different about pitchers with a higher usage of cutters. Despite there being little variation in exit velocity and hard hit percentage, the cutterballers were the only group with unique results. This was not only true for some of the quality of contact variables; in general, cutterballers performed better in terms of ERA than other pitchers. Future research might examine pitchers with cutters further and investigate differences between other types of pitchers and benefits of cutter usage.

One flaw of this analysis is that the Statcast custom leaderboards categorizes pitches into 8 different pitch types. This is obviously a problem when Yu Darvish throws 10 different pitches and some pitchers blur the lines between pitch types (Petriello 2019). Darvish throws two different types of cutters and a slider that isn't that different from his cutters. Miscategorization of pitches could influence these results. But as pitch tracking systems become more sophisticated in the future, research might become less affected by this.

Another problem with this analysis is a few important variables were not included. Average velocity wasn't included because it was only available per individual pitch type in the dataset and not all pitchers throw each type of pitch. As discussed in the introduction, velocity has been found to be a factor in quality of contact. Another potential important factor that wasn't

included in the data source is pitch tunneling, which is the idea of a pitcher throwing different types of pitches that have very similar flight paths until after the batter has to decide if they're swinging or not (Pavlidis, Judge, and Long 2017). This analysis didn't include any metrics regarding deception or release point. Furthermore, the sequencing of pitch types wasn't considered, but could be important. Other important factors could include ballpark effects, pitcher handedness, if the pitcher is a starter or a reliever, and how the ball is playing that year.

Additionally, this analysis looked at grouping individual pitcher seasons, but it may make sense to also examine individual pitcher months due to frequent adjustments. Future research could also focus on individual batted ball events rather than the performance of specific pitchers. Investigating all these overlooked factors could provide avenues for future research. Other types of analysis, such as factor analysis, might be able to provide more understanding about quality of contact.

References

Arthur, R. (2015, May 22). A Baseball's Exit Velocity Is Five Parts Hitter, One Part Pitcher. Retrieved October 26, 2019, from FiveThirtyEight website: https://fivethirtyeight.com/features/a-baseballs-exit-velocity-is-five-parts-hitter-one-part-pitcher/

Carleton, R. A. (2016, April 19). Baseball Therapy: The One About Exit Velocity - Baseball Prospectus. Retrieved October 26, 2019, from Baseball Prospectus website: https://www.baseballprospectus.com/news/article/28956/baseball-therapy-the-one-about-exit-velocity/

Casella, P. (2015, April 24). Statcast Primer: Baseball Will Never Be The Same. Retrieved April 12, 2020, from MLB website: https://www.mlb.com/news/statcast-primer-baseball-will-never-be-the-same/c-119234412

Clemens, B. (2019, May 31). Gregory Groundball vs. Marty McFly: Who Allows More Big Innings? Retrieved October 26, 2019, from FanGraphs Baseball website: https://blogs.fangraphs.com/gregory-groundball-vs-marty-mcfly-who-allows-more-big-innings/

Day, Z. (2013, December 5). Measuring Pitching with TrackMan: The Secrets of Fastball Spin – Baseball Prospectus. Retrieved October 26, 2019, from Baseball Prospectus website: https://www.baseballprospectus.com/news/article/22362/measuring-pitching-with-trackman-the-secrets-of-fastball-spin/

Fast, M. (2011, November 16). Spinning Yarn: Who Controls How Hard the Ball is Hit? – Baseball Prospectus. Retrieved October 26, 2019, from Baseball Prospectus website: https://www.baseballprospectus.com/news/article/15532/spinning-yarn-who-controls-how-hard-the-ball-is-hit/

McCracken, V. (2001, January 23). Pitching and Defense: How Much Control Do Hurlers Have? - Baseball Prospectus. Retrieved October 26, 2019, from Baseball Prospectus website: https://www.baseballprospectus.com/news/article/878/pitching-and-defense-how-much-control-do-hurlers-have/

O'Connell, M., & Marsh, J. (2016, November 17). Spin Rate: What We Know Now. Retrieved October 26, 2019, from Driveline Baseball website: https://www.drivelinebaseball.com/2016/11/spin-rate-what-we- know-now/

Pavlidis, H., Judge J., & Long J (2017, January 23). Prospectus Feature: Command and Control. Retrieved April 10th, 2020 from Baseball Prospectus website: https://www.baseballprospectus.com/news/article/31022/prospectus-feature-command-and-control/

Petriello, M (2019, September 4). Not a typo: Yu Darvish throws 10 pitches. Retrieved April 14th, 2020 from MLB website: https://www.mlb.com/news/cubs-yu-darvish-throws-10-different-pitch-types

Sarris, E. (2019, August 30). Is a pitcher's ground-ball rate more important in today's game? Retrieved October 26, 2019, from The Athletic website: https://theathletic.com/1171344/2019/08/30/is-a-pitchers-ground-ball-rate-more-important-in-todays-game/

Statcast Leaderboard. (2019). Retrieved October 26, 2019, from baseballsavant.com website:

https://baseballsavant.mlb.com/statcast_leaderboard

Tippet, T. (2003, July 21). Can pitchers prevent hits on balls in play? Retrieved October 26, 2019, from Diamond Mind Baseball website: http://web.archive.org/web/20040208104252/http:/www.diamond-mind.com/articles/ipavg2.htm

Zimmerman, J. (2015, November 30). kwERA: The Starting Point for Pitcher Evaluations. Retrieved October 26, 2019, from The Hardball Times website: https://tht.fangraphs.com/kwera-the-starting-point-for-pitcher-evaluations/

Zimmerman, J. (2016, May 6). Can Pitchers Prevent Solid Contact? Retrieved October 26, 2019, from The Hardball Times website: https://tht.fangraphs.com/can-pitchers-prevent-solid-contact/