Reclassifying Relief Pitchers

Literature Review

History of Baseball and Relief Pitching

Rules, structure, and team strategy of baseball that have progressed throughout time. Presented by Woltring *et al*[1], many of the eras left an impact on the utilization of relief pitchers. This thesis will be covering some of the eras through the course of baseball's foundation. Table 1 shows all the eras with their respective years of existence and changes.

Table 1: *Baseball Eras*

Era	Years	Notes
Pre-Dead Ball	1839 – 1900	1) Origin of the game
Dead Ball	1901 – 1919	1) home plate is changed from a square to base creatin larger strik 2) Foul balls begin to count as strikes
Live Ball	1920 – 1941	1) American League was created 2) Spitball and Emery ball were removed from the game 3) Baseball was reconstructed allowing more success for hitters
Integration	1942 – 1960	1) Many players were replaced because of service time in WW2 2) Integration of all athletes regardless of skin color
Expansion	1961 – 1976	1) Expansion in teams from 16 to 24 2) Strike zone was redefined: shoulders – bottom of the knee 3) Slope of mound decreased from 15 to 10 inches above home plate (1969) 4) Creation of Designated Hitter in American League (1973)
Free Agency	1977 – 1993	1) Free agency caused a restructure in team strategy 2) Player salaries drastically increased
Steroid	1994 – 2005	1) Many star players were increasing power by performance-enhancing drugs (PED's) which were banned by MLB
Post-Steroid	2006 – *Present*	1) Punishment for steroids became significantly more threatening 2) Advanced research on pitching metrics resurrecting pitching dominance

As rules began to change and strategy shifted, so did the philosophy of relievers. During the Pre-Dead Ball Era, the concept of a relief pitcher was admirable. In most cases, starting pitchers threw the whole game. Little to no research on the safety of a pitcher's arm or recovery resulted in ball clubs believing that using their best pitcher for the entirety of the game was the best way to be successful. Per Chris Bodig *et al*[2], table 2 shows how often starters threw for complete games during different eras.

Table 2: *Totals Among Pitchers Per Decade*

- **% G as CG:** Percentage of complete games over the decade
- **% G as RP:** Percentage of games using a reliever over the decade
- **GR 50 +:** Number of pitchers who threw at least 50 games in relief over the decade
- **GR 100+:** Number of pitchers who threw at least 100 games in relief over the decade

Decade	% G as CG	% G as RP	GR 50+	GR 100+
1871-1879	92.0%	7.6%	0	0
1880-1889	93.9%	6.0%	0	0
1890-1899	85.4%	13.5%	1	0
1900-1909	79.0%	19.4%	10	0
1910-1919	56.9%	36.7%	109	6

As we can see, the Pre-Dead ball Era almost exclusively relied on their starters to throw the entirety of games. From 1871-1899, there was only one pitcher in all of baseball who pitched in relief over 50 times over a decade. Per Chris Bodig that pitched resulted in Hall of Famer Ty Cobb. Bodig explains a lot about how relievers were looked at during this period. In the rare cases that a reliever needed to be used, the ball club's philosophy was to throw their best pitcher available. There were no "designated relievers." There were starting pitchers and starters who did not start that day who were available to throw later in the game if need be. Ty Cobb was one of the best pitchers of his era so it makes sense that he was the only reliever with over 50 appearances. The Cleveland Spiders relied on the best arm in the game to come in whenever they needed someone else to throw. At the same time (1871-1910), there was only one pitcher who had more relief appearances than starts; "Doc" James Otis Crandall. He was nicknamed "Doc" because he excelled in late-game situations to save his team like a doctor would with a patient. Even though he started in roughly 40% of his outings, he was the first sign baseball had seen of a true reliever. The strategy of pitcher usage has changed drastically compared to the modern-day. The beginning of the Dead Ball Era was when we began to see a relative uptick in the utilization of relievers.

Once the 1920's began, a shift in advantage from the defense to offense materialized. Spitballs were removed from the game. Baseballs were replaced mid-game to ensure that the ball was always white so players could see it. Most importantly, the baseball was redesigned with a “new cushioned cork center” in the middle to allow for more offense production. With the advantage in the hands of the offense, it was time to reconstruct how baseball used their pitchers. Frederick “Firpo” Marberry was the first star reliever. Coming out of the bullpen 114 times and only making 5 starts between 1925-1926, he was coined a true arm that excelled in getting saves (even though the term was not used until 1969). Soon followed many pitchers who were called upon in late-game situations. These pioneers were eventually called “closers.” Some notable starting pitchers who were called upon to close are listed below

Table 3: *Starting Pitchers Used as Closers (Pre-Dead Ball Era – Beginning of Integration Era [1945])*

- **Orange:** Hall of Fame Inductee

Name	W	L	SV	G	GS	IP	K/9	BB/9	HR/9	BABIP	ERA	FIP	WAR
Firpo Marberry	148	88	101	551	187	2067.1	3.58	2.99	0.42	0.274	3.63	3.92	22.9
Johnny Murphy	89	51	97	356	40	945.1	3.33	3.78	0.45	0.264	3.54	4.12	4.4
Clint Brown	89	93	64	434	130	1485.2	2.48	2.23	0.51	0.295	4.26	4.03	18.9
Joe Heving	76	48	63	430	40	1038.2	3.72	3.29	0.55	0.291	3.9	4.01	6
Jack Quinn	247	218	57	756	444	3920.1	3.05	1.97	0.23	0.291	3.29	3.15	61.2
Lefty Grove	300	141	55	616	457	3940.2	5.18	2.71	0.37	0.284	3.06	3.36	88.8
Waite Hoyt	237	182	52	674	423	3762.1	2.88	2.4	0.37	0.283	3.59	3.76	48.9
Mordecai Brown	239	130	49	481	332	3172.1	3.9	1.91	0.12	0.26	2.06	2.36	49.3
Ace Adams	41	32	49	299	7	550	2.75	3.65	0.39	0.262	3.4	3.72	0.7

As the role of relief pitchers grew, so did their recognition by the media and fans. From Chris Bodig *et al*[3], Page was the first “closer” to earn a relevant amount of MVP votes. Although he would be considered today as the closer for the New York Yankees at the time, he was not utilized like closers are today. For example, in game 7 of the World Series, Page came in to throw innings 5-9 to secure the championship for his team. The relievers were stretched out over a much longer period of innings per average appearance. However, this was a big step in the right direction. Jim Konstanty became the first reliever to win the MVP in 1950. In the years to come, we see an emergent of relief pitching stars who specialized in preventing offensive production better than ever before. Some Hall of Fame relievers that occurred over the next 30 years are Goose Gossage, Lee Smith, Dennis Eckersley, Rollie Fingers, Sparky Lyle, Al Hrabosky and Steve Bedrosian.

As baseball continued to grow until the state of the current game, reliever utilization did as well. The more time that has passed, the less likely a reliever was expected to throw for an extended amount of innings per appearance. Table 4 shows how closers are on “shorter leashes” meaning they are not throwing as many innings per outing.

Table 4: *Average Innings Pitched for Top 20 Closers (ranked by saves)*

Years	Average IP per save		% of Saves 1 IP or less	% of Saves 2 IP or more
1970-74	1.62		43%	24%
1975-79	1.68		42%	28%
1980-84	1.62		39%	26%
1985-89	1.35		52%	21%
1990-94	1.13		72%	9%
1995-99	1.03		86%	2.3%
2000-04	1.03		88%	2.2%
2005-09	1.01		92%	0.9%
2010-14	0.99		96%	0.2%

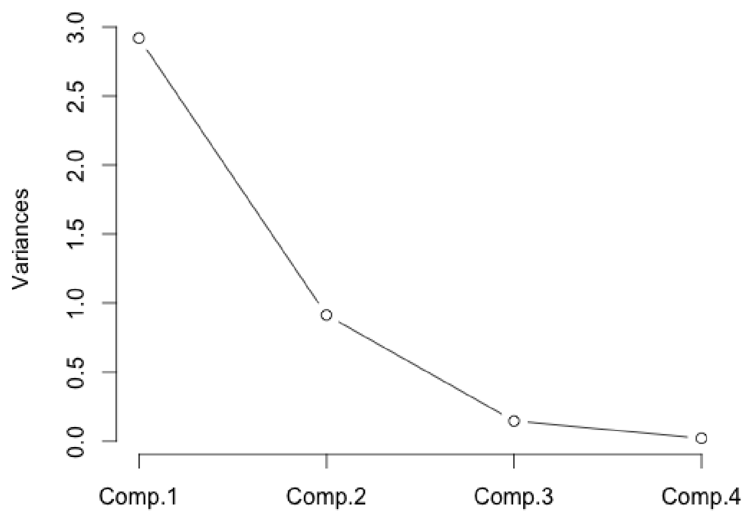
As we can see, closers do no throw nearly the amount of batters they used to in the 70's. This shift is the last major change in reliever usage until the introduction of the Opener in 2018 by the Tampa Bay Rays.

Clustering

There are many forms in statistical data science to analyze data successfully. One of the more robust methods is clustering. According to Kumar *et al*[4], Cluster analysis divides data into groups (clusters) that are meaningful, useful, or both. It can also be used as a "starting point" for further analysis; primarily being data summarization. The greater the homogeneity within a group and the larger the difference between groups, the more distinct clustering results. With the proper data, one can learn much more about the similarities differences in the dataset based off patterns discovered by machine learning. There are multiple forms of clustering data scientist use depending on what they are trying to achieve. For now, we will focus on K-Means Clustering Analysis.

K-Means Clustering is one of the oldest and most widely used clustering algorithms. Essentially, K-Means self-determines a best-fit centroid by finding the mean of a group of points in a continuous *n*-dimensional space. For successful analysis, we must determine an initial number of centroids, or groups, which we will represent as *K.* One method to determine the optimal number for K based off your data is the "Elbow Method." PlanSpace *et al*[5] explains how The Elbow Plot determines the ideal number of centroids by graphing the variance within the cluster analysis. Once the graph shows a strong bend (like an elbow) we can determine the ideal number of centroids needed for a dataset. An example from PlanSpace is shown below using the iris dataset given in RStudio.

Figure 1: *Elbow Plot (Iris Dataset)*



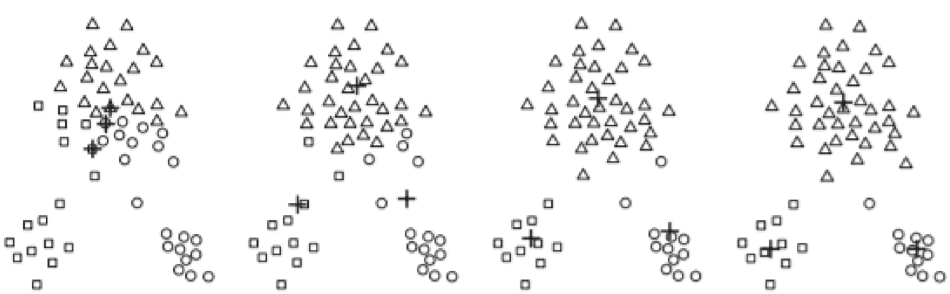
Comp. 2 serves as the optimal iterations of cluster for the Iris dataset using the Elbow Method.

Now that we have calculated K, we can begin the clustering process. Figure 2 explains the process thoroughly. As more data is included into our clustering algorithm, the centroids will continue to move closer to their target mean of each group until they cannot move any further. This proves is shown in Figure 3.

Figure 2: *Basic K-Means Algorithm*

```
Select K points as initial centroids.
repeat
    Form K clusters by assigning each point to its closest centroid.
    Recompute the centroid of each cluster.
until Centroids do not change.
```

Figure 3: *K-Means Algorithm in Sample Data (K =3)*



(a) Iteration 1. (b) Iteration 2. (c) Iteration 3. (d) Iteration 4.

One question that has yet to be answered is the following: “How do we determine what centroid is closest to each point?” Calculating Euclidean distance is used to answer this question. Typically, programing languages like RStudio and Python calculate this for you, but we can calculate rather simply. By using the sum of square error (SSE), we can calculate the error of each data point (or the Euclidean distance to the centroid) and sum all the squared errors.

Figure 4: *SSE Formula*

Symbol	Description
$\mathbf{x}$	An object.
C_i	The i^{th} cluster.
$\mathbf{c}_i$	The centroid of cluster C_i.
$\mathbf{c}$	The centroid of all points.
m_i	The number of objects in the i^{th} cluster.
m	The number of objects in the data set.
K	The number of clusters.

$$\text{SSE} = \sum_{i=1}^{K} \sum_{\mathbf{x} \in C_i} dist(\mathbf{c}_i, \mathbf{x})^2$$

Using the notation above, *dist* is the “standard Euclidean distance between two objects in Euclidean space” (Kumar *et al*[4]*).* One last issue remains in K-Means form of cluster analysis. When random initialization of centroids is used, multiple runs of K-Means will produce different SSE results. There are many ways to work around this. The most popular is to run multiple runs and select the set of clusters with the Minimum SSE. This way, we know our centroids are most closely fit to their cluster. Once we have our minimal SSE and our ideal clusters, we have completed the statistical analysis but there is ore to come Next, we must make sense of the clusters in order to create an accurate conclusion.

Clustering Effects in Baseball

Using the machine learning tool listed above, we can draw conclusions in baseball that can help advance the game in the future. Gerlica *et al*[6] tested the use of clustering on outfield shifts. Their goal was to find a more accurate way for the Air Force College baseball team to determine where to shift their players on defense. Originally, the coaching staff would look at scouting reports of their opponent’s offense and adjust accordingly based on previous results. The groups plan was to evaluate catch probability of all flyballs in the outfield to move their outfielders in positions to maximize the likeliness of any given flyball being caught regardless of who was up to bat. Using past locational data with the position of Air Force’s outfielders when each flyball was caught, the group was able to use K-Means cluster analysis to determine the ideal position of each outfielder. In Figure 5, the team plots the locational data for each ball caught and determines the means of each cluster to create their *K* centroids. Once the centroids are created, the testing can begin, and the results are shown in Figure 6.

Figure 5: *Determining Centroid Location on Previous Ball-in-Play- Results*

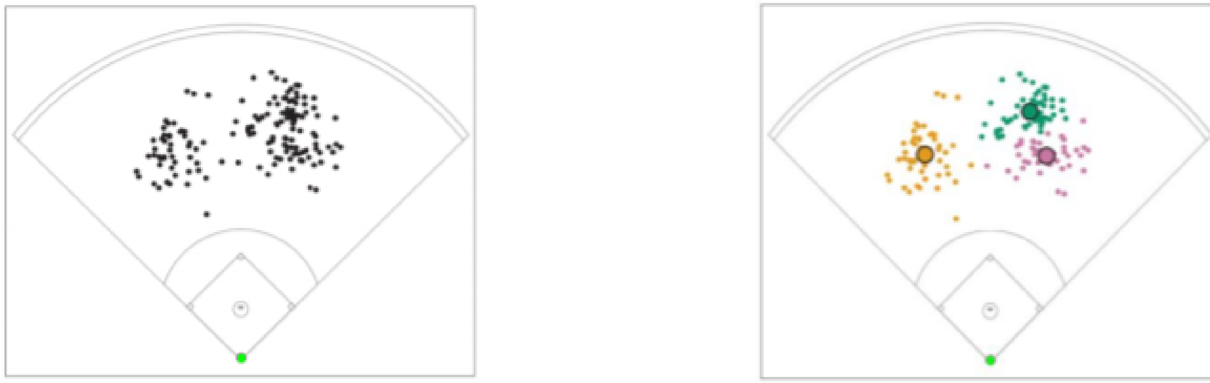
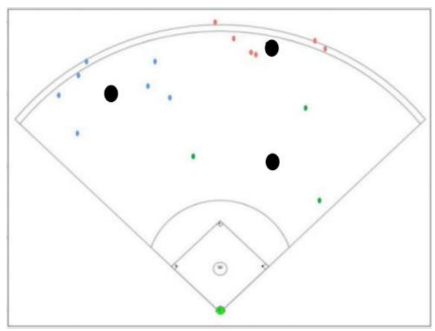


Figure 6: *Adjusted Position of Outfielders using K-means Clustering Techniques*



The group concluded that the adjusted positioning of outfielders would result in 7.1% increase of outs created by the outfielders. The group concluded that this position shift in outfielders would be relevant enough to gain another 1-2 wins from the previous season. Two wins would have been enough for Air Force to take first place in their conference and ultimately a bid to the college tournament. Although Air Force did not end up using the findings of their students, the team was confident that their findings would lead to greater success on defense, saving runs and helping the team generate more wins. This project proved that cluster analysis can show findings to help teams be more successful if teams decided to act upon the findings.

Using Clustering and Other Analysis on Relievers

Now that we have discussed the history and utilization of relievers along with learning about the powers of clustering analysis, we can combine both and generate findings. Looking at the current roles of relievers, I realized there are strengths each reliever has that can be used more effectively in games to generate more success. One major statistic that needs to be looked at more carefully is groundball percentage. Sarris *et al*[7] *goes* into depth about the effectiveness of groundball pitchers in today's game. Sarris explains how as the game has progressed, the ground ball pitcher has become more and more important for a staff. The reason being is simple. As the game has progressed, the long ball has become significantly more valuable. Currently, certain teams can get away with a win from only 4 hits, if some of those landed over the outfield wall. We see successful teams leading the league in power metrics. Therefore, having pitchers who can keep the ball on the ground explode in value. These are the pitchers who can counter the way offensive lineups in baseball are shifting. Although this is information, I believed prior, this article solidified my theory. This is important for my project because it stresses the strength of the ability to get groundballs. Marcou *et al*[8] performs cluster analysis on all pitchers to see if he can find patterns in pitchers that do have high ground ball rates. Although he did not come up with any relevant conclusions, it gave me confidence that clustering on pitchers can generate some sort of conclusion.

Zimmerman *et al*[9] questions if pitchers can prevent solid contact. Zimmerman uses stats from Baseball Info Solutions, Statcast, and FanGraphs to show real examples of players' performance to get weak contact. He uses a main example of Dallas Keuchel, who at the time, had a sharp decrease in soft contact from his most recent season to his previous ones prior. Zimmerman attempts to uses analyze his stats to come up with why that happened. He notices that a change in average velocity could be part of the reason he was giving up harder contact. I think we also have to take into consideration how Keuchel is facing the same teams over again, and the more he faces the same batters, the more likely they will get an understanding of what he throws and when he varies his pitches. This will help a batter's success in making hard contact. This article made me realize that I needed to come up with a factor for pitchers that looks at how long he has been throwing against the same teams. I feel that would be useful, however I don't know if it is feasible. Regardless, accounting for hard hit percentage in my analysis is crucial.

Cermark *et al*[10] determines the difference between percentage and "per-nine" statistics. The primary example used is Strikeout Percentage (K%), and Strikeout Per 9 Innings (K/9). Before looking into Cermark's research, I underestimated the difference between the metrics. He concluded that "per-nine" numbers don't tell the whole story. For example, a pitcher with a 12 K/9 does not include how often they give up hits. A pitcher with a high K% proves that the pitcher gets strikeouts often, because the stat accounts for all batters the pitcher faces. K/9 only accounts for the outs a pitcher creates. I must use K% for my cluster analysis to fully determine the effectiveness of a high-strikeout pitcher. If there is an option for a percentage stat over a "per nine" I will use the percentage statistic for my research.

References

1. Woltring, M. T., Rost, J. K., & Jubenville, C. B. (2018, October 25). Examining Perceptions of Baseball's Eras: A Statistical Comparison. Retrieved November 01, 2020, from https://thesportjournal.org/article/examining-perceptions-of-baseballs-eras/
2. Bodig, C. (2017, October 23). The History of Relief Pitching Part One: 1871-1945. Retrieved November 01, 2020, from https://www.cooperstowncred.com/the-history-of-relief-part-one-1871-1945/
3. Bodig, C. (2017, October 23). The History of Relief Pitching Part Two: 1946-1968. Retrieved November 02, 2020, from https://www.cooperstowncred.com/the-history-of-relief-pitching-part-two-1946-1968/
4. ________________________________
5. Schumaker, AJ. *PCA, 3D Visualization, and Clustering in R*, 3 Feb. 2013, planspace.org/2013/02/03/pca-3d-visualization-and-clustering-in-r/.
6. Gerlica, Jeffrey, et al. "Quantifying the Outfield Shift Using K-Means Clustering." *ieworldconference.org*, 30 Apr. 2020, www.ieworldconference.org/content/WP2020/Papers/GDRKMCC_20_85.pdf.
7. Sarris, Eno. "Is a Pitcher's Ground-Ball Rate More Important in Today's Game?" *The Athletic*, The Athletic, 29 Aug. 2019, theathletic.com/1171344/2019/08/30/is-a-pitchers-ground-ball-rate-more-important-in-todays-game/.
8. Marcou, Charlie. "Investigating Major League Baseball Pitchers and Quality of Contact through Cluster Analysis." *ScholarWorks@GVSU*, 30 Apr. 2020, scholarworks.gvsu.edu/honorsprojects/765/.
9. Zimmerman, Jeff. "Can Pitchers Prevent Solid Contact?" *The Hardball Times*, tht.fangraphs.com/can-pitchers-prevent-solid-contact/.
10. Cermak, Ben. "K% Vs. K/9: Finding an Effective Metric for Strikeout Rate." *Statliners*, FanSided, 10 Nov. 2014, statliners.com/2014/11/10/k-vs-k9-finding-effective-metric-strikeout-rate/.
11.