

빅데이터 베이스볼: 데이터로 알아보는 야구

6강: Visualizing Baseball

장원철

서울대학교 통계학과

wcjang@snu.ac.kr

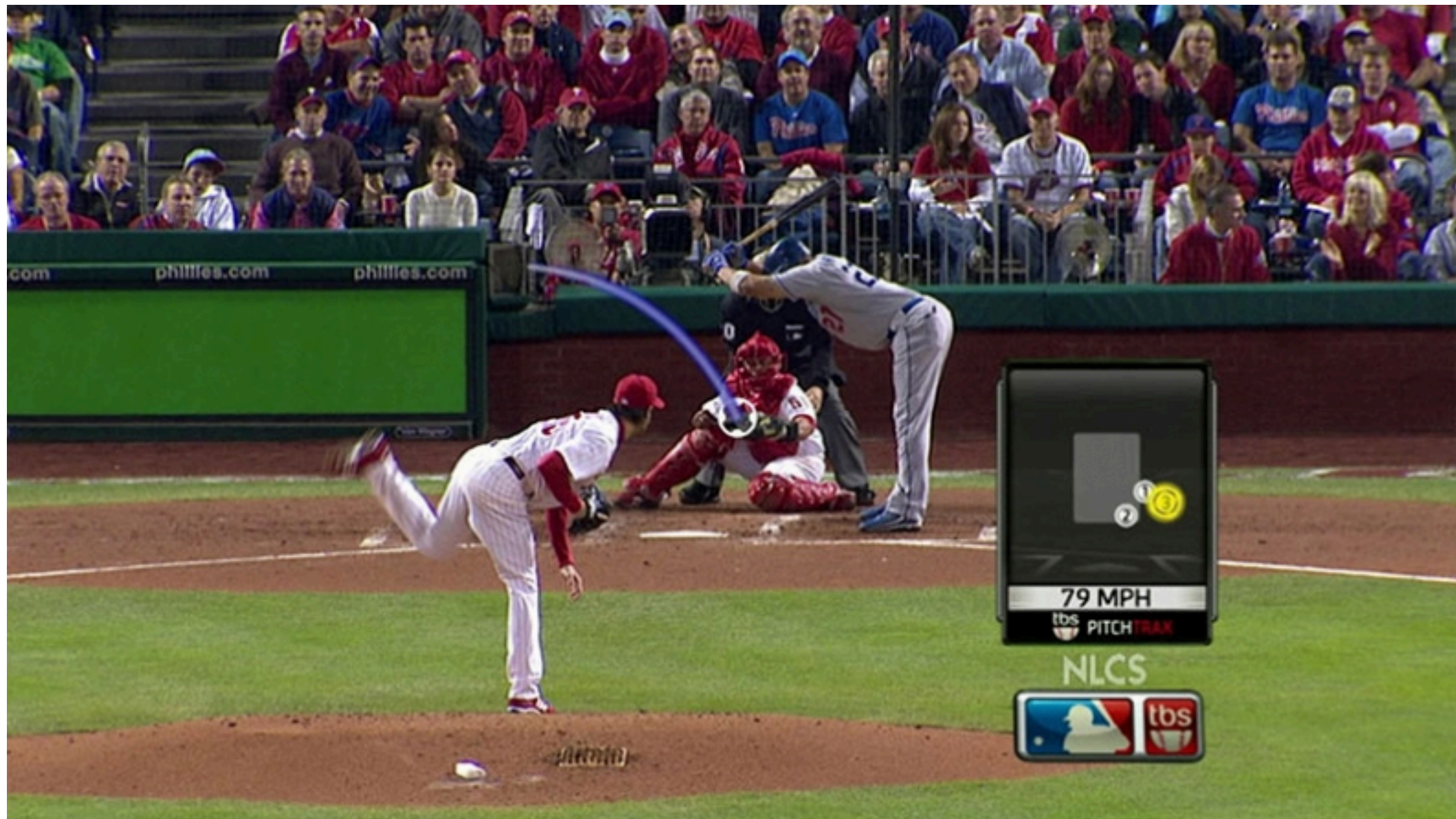
SportsVision

- 1996년 Stan Honey가 아이스하키 경기 TV 중계시 시청자들이 하키펍을 종종 놓친다는 사실에 착안해서 보다 쉽게 시청자들이 하키펍을 쉽게 찾을 수 있는 시각적인 방법을 제시 <https://www.youtube.com/watch?v=grOttSHuuzE>
- 이후 미식축구 TV 중계에서 first down line을 노란색으로 표시하는 방식을 도입하여 대성공을 거둠



Sportvision

- PITCHf/x system을 개발하여 MLB에 공급



MLB Plus

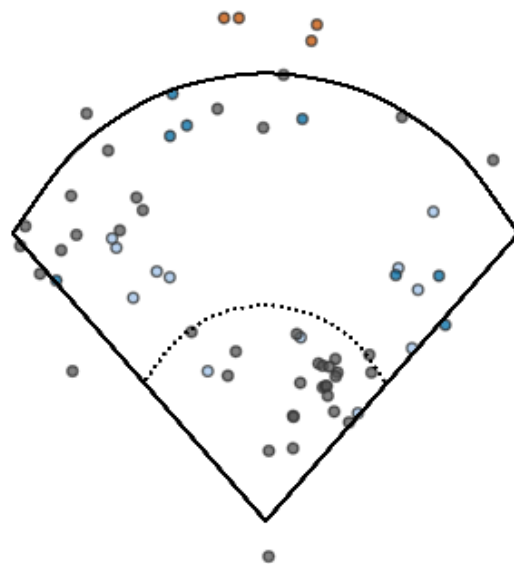
- Statcast를 바탕으로 data-driven online 야구중계



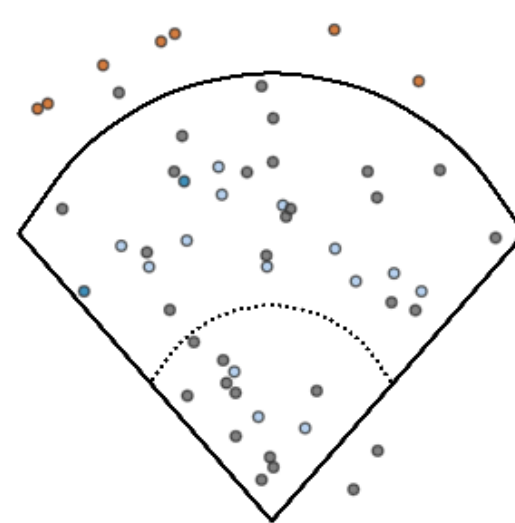
Shiny App

- The interactive Spray Chart Tool

Bryce Harper



Mike Trout



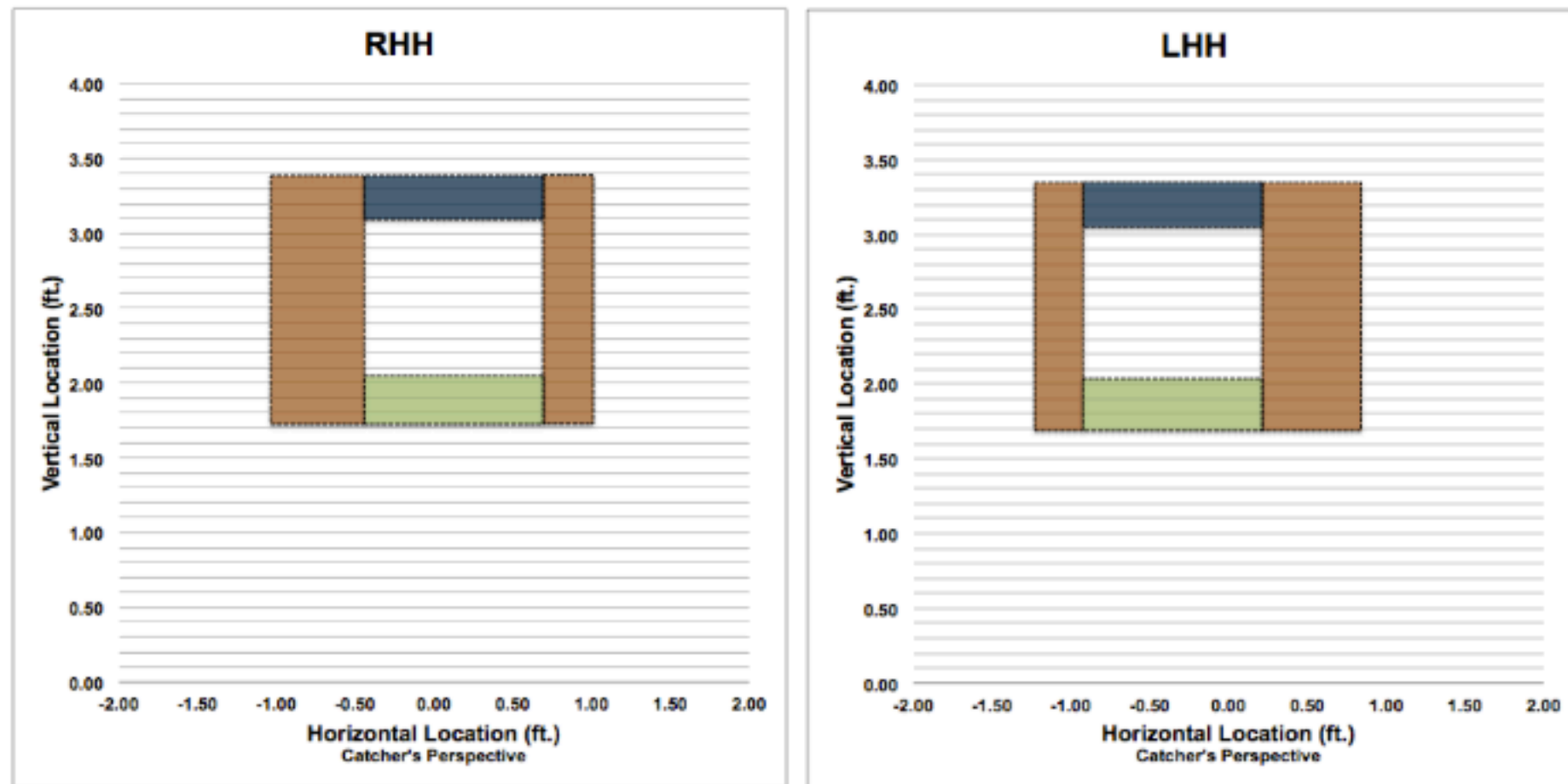
Outcomes • 1B • 2B • HR • Out

Designed and maintained by Bill Petti
billpetti.github.io
Data courtesy of MLBAM/baseballsavant.com

	Name	Batted Balls	Barrel %	Average Launch Angle (degrees)	Average Launch Speed (mph)	Expected wOBA on Contact	Actual wOBA on Contact	SLG on Contact
1	Bryce Harper	65	0.077	13.5	91.3	0.501	0.471	0.677
2	Mike Trout	55	0.164	19.6	88.8	0.525	0.552	0.873

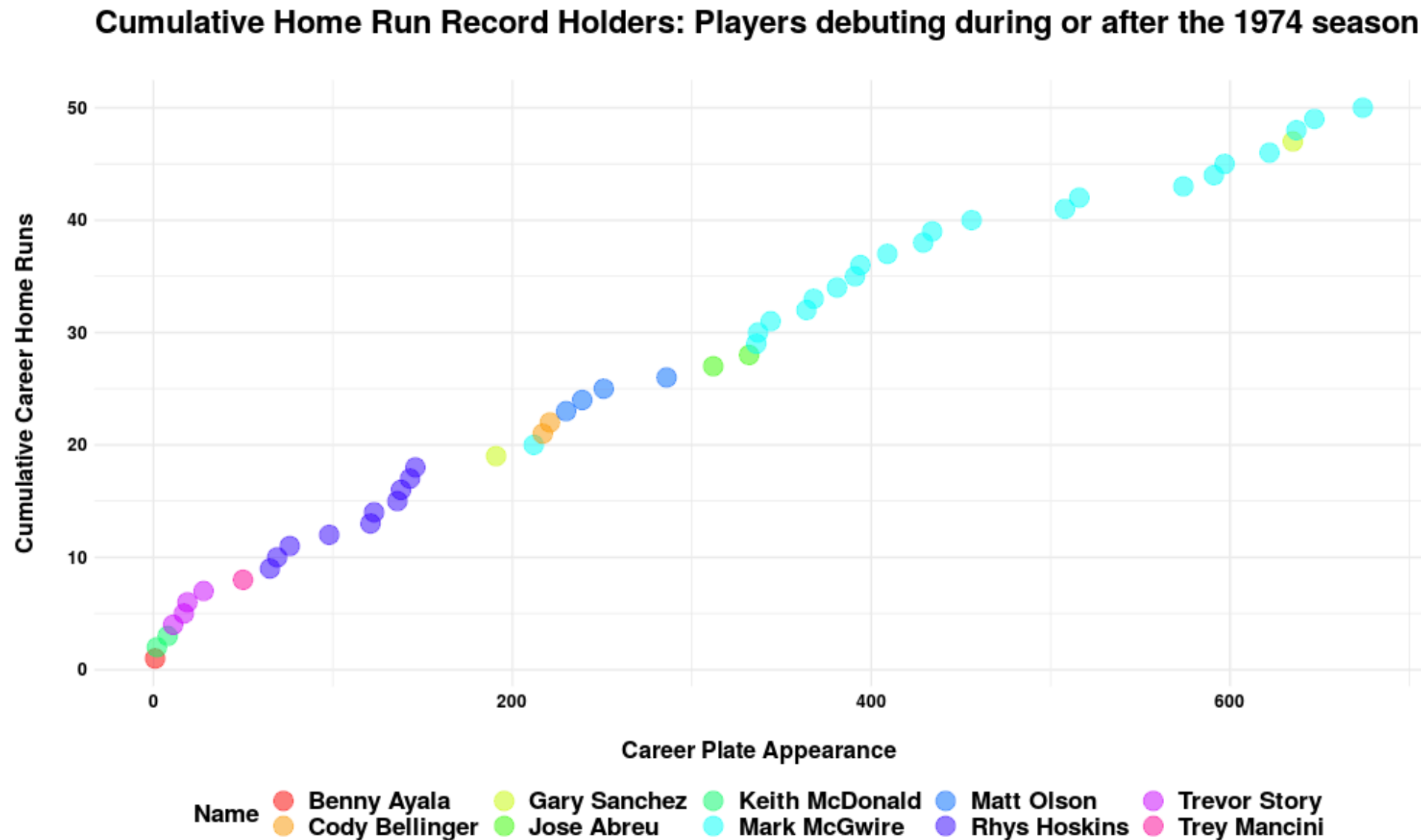
Shiny App

- 타자, 투수, 팀별 스트라이크 존의 위치별 투구수



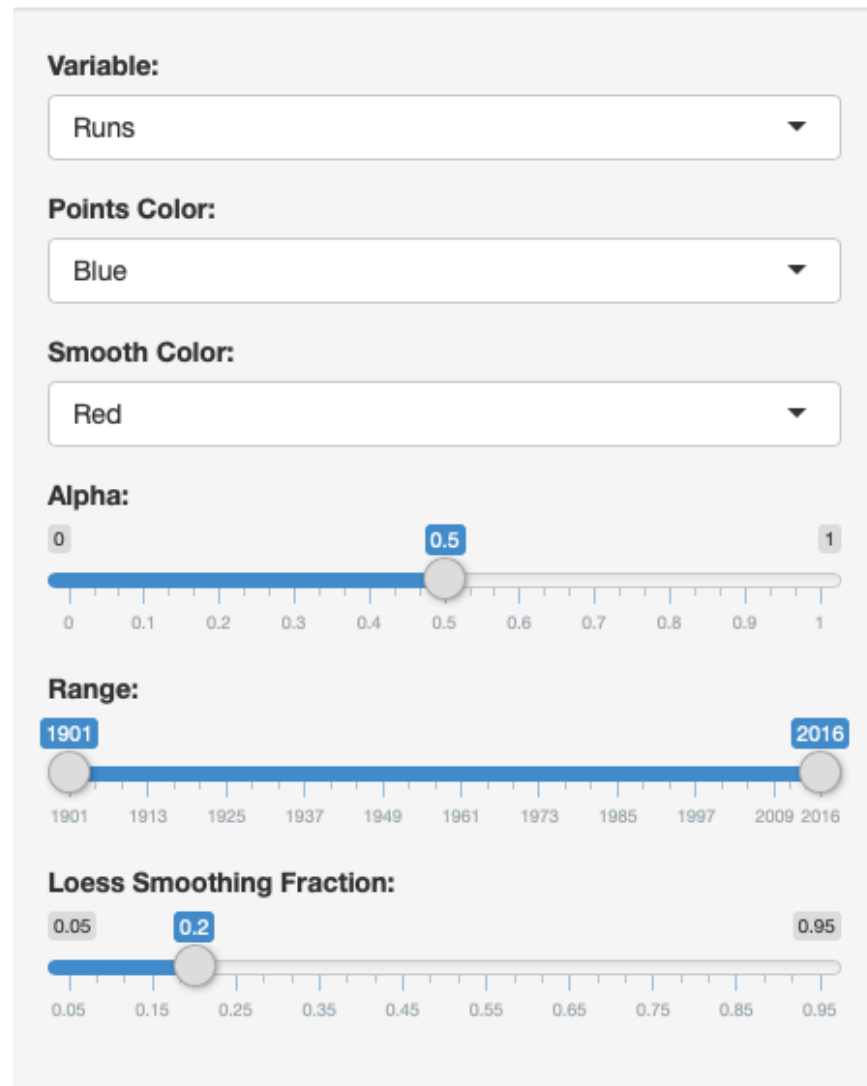
Shiny App

- 홈런당 최소타수를 기록한 선수

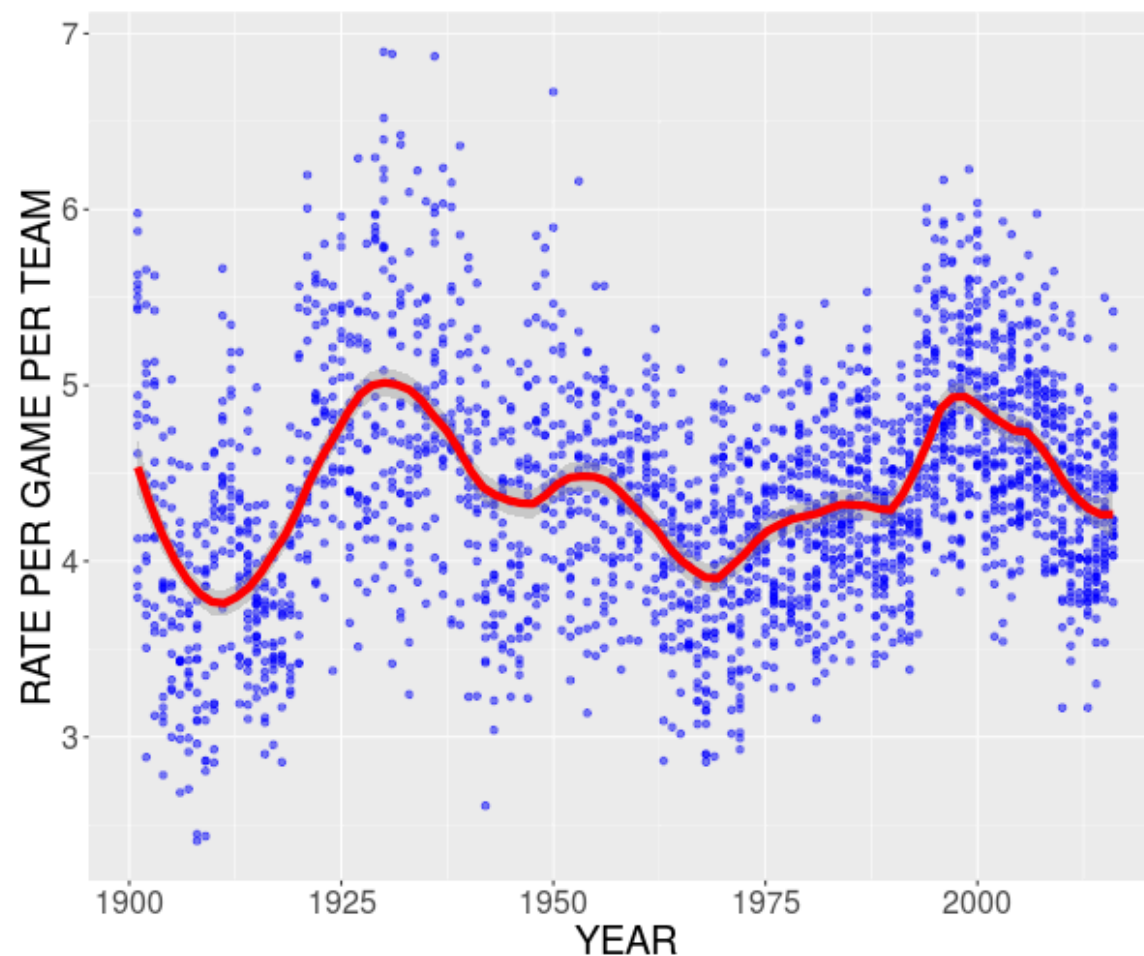


Shiny App

- Offensive Production in Baseball History

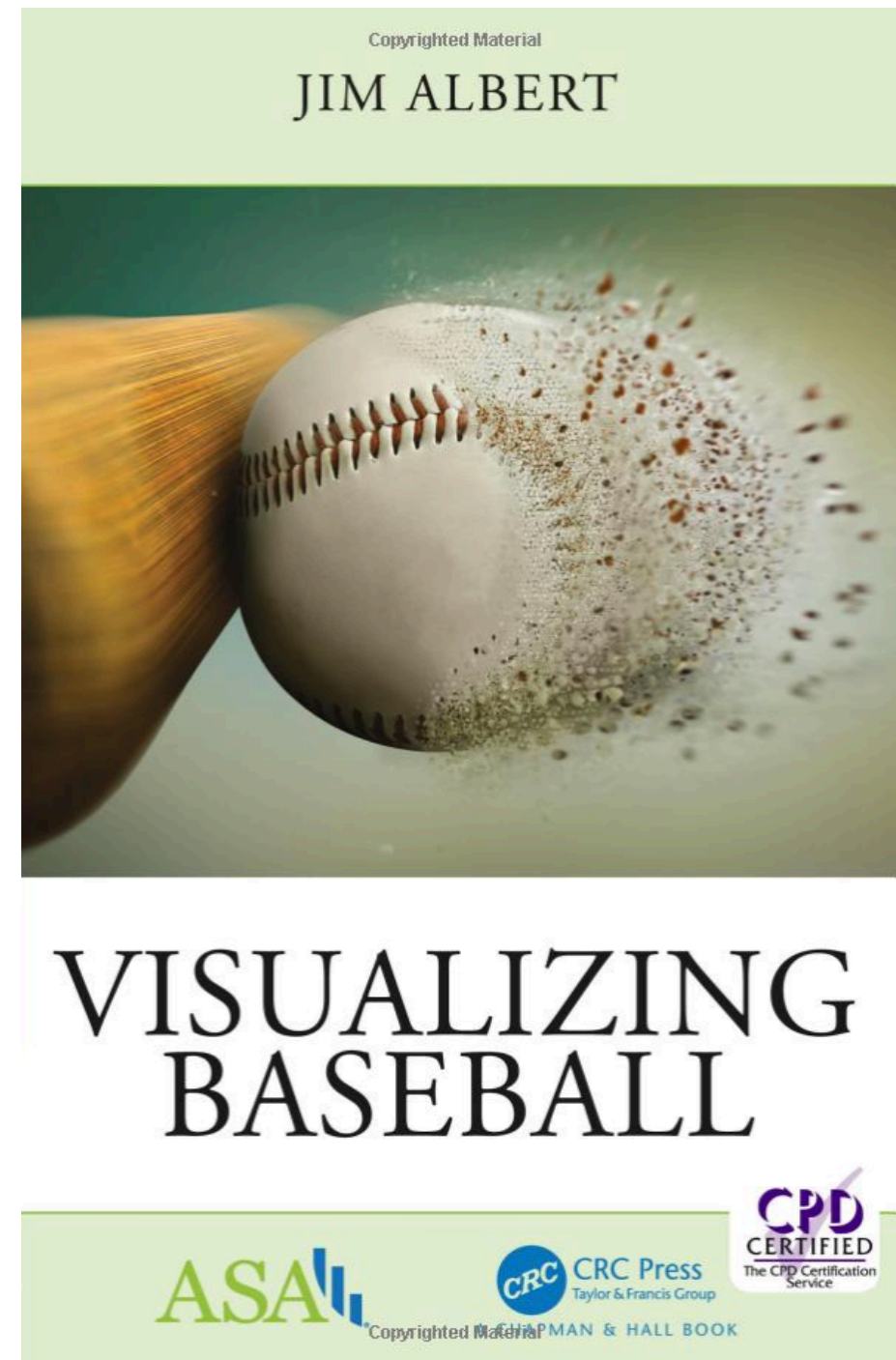


Variable: R Per Game Per Team: 1901 to 2016

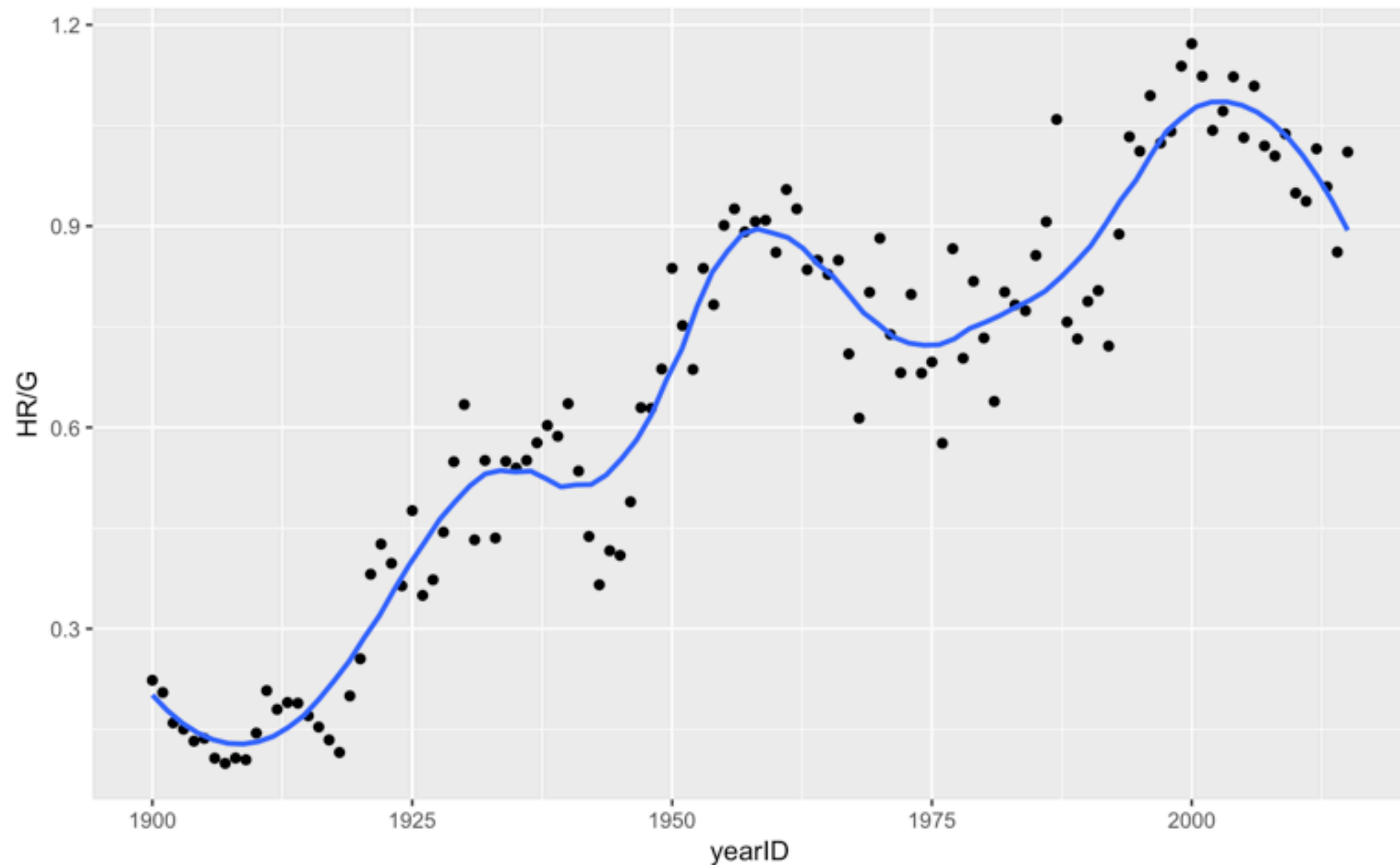


Visualizing Baseball

- Jim Albert (2017)
Visualizing Baseball
- R ggplot2 package 사용
- Codes는 저자의 깃허브 페이지에서 제공 <https://bayesball.github.io/VB/>

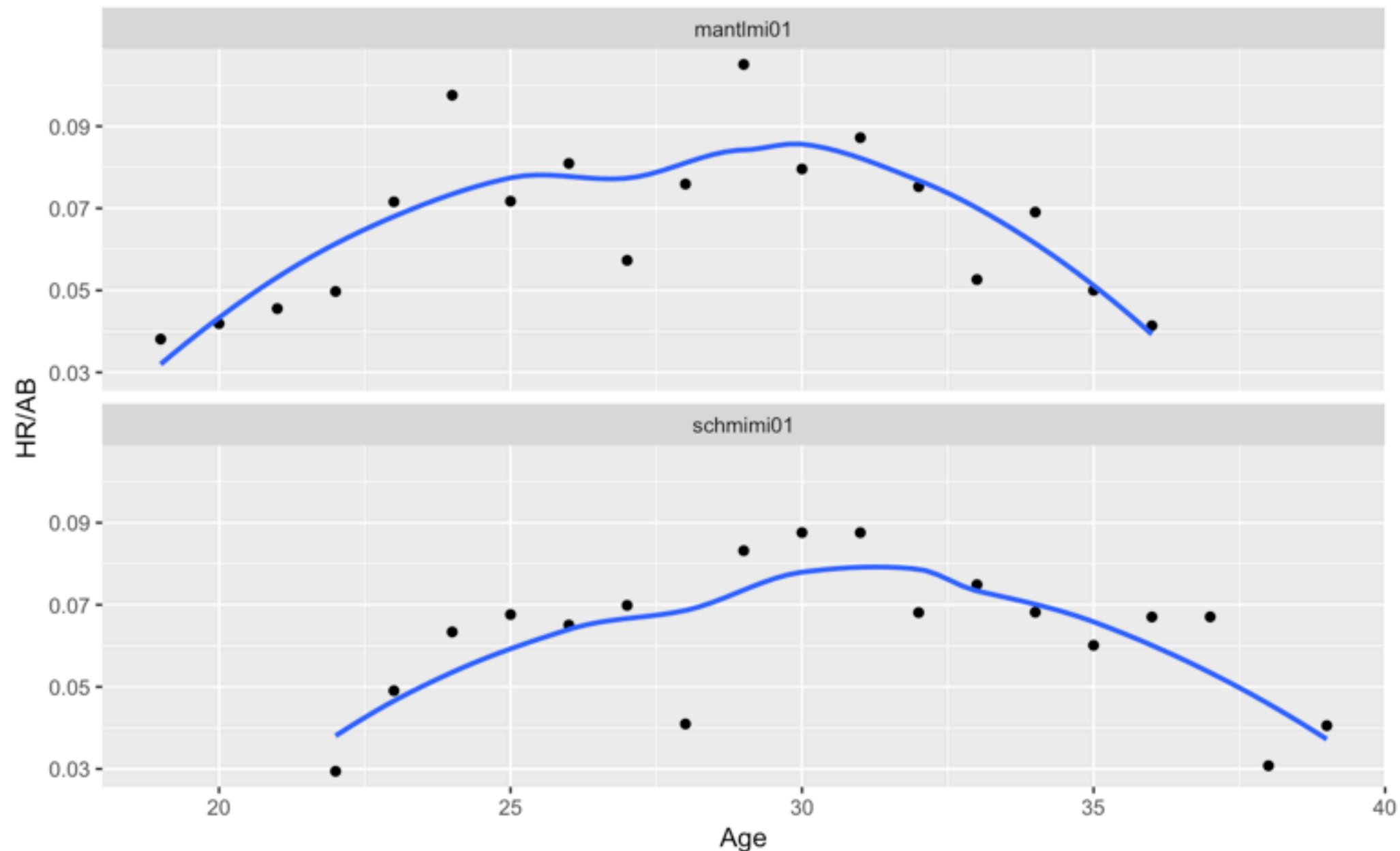


History of Home Run Hitting

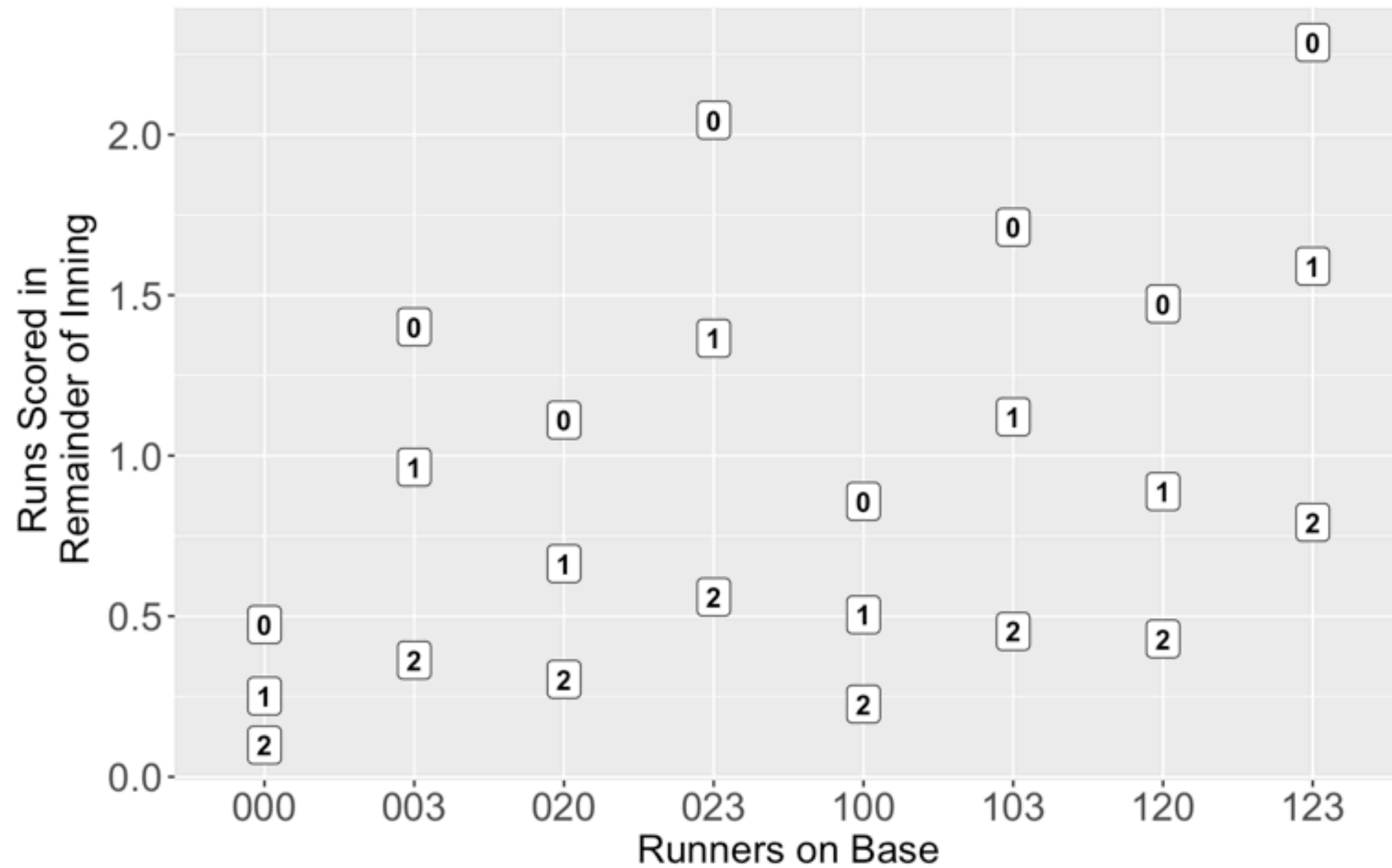


Comparing Trajectories

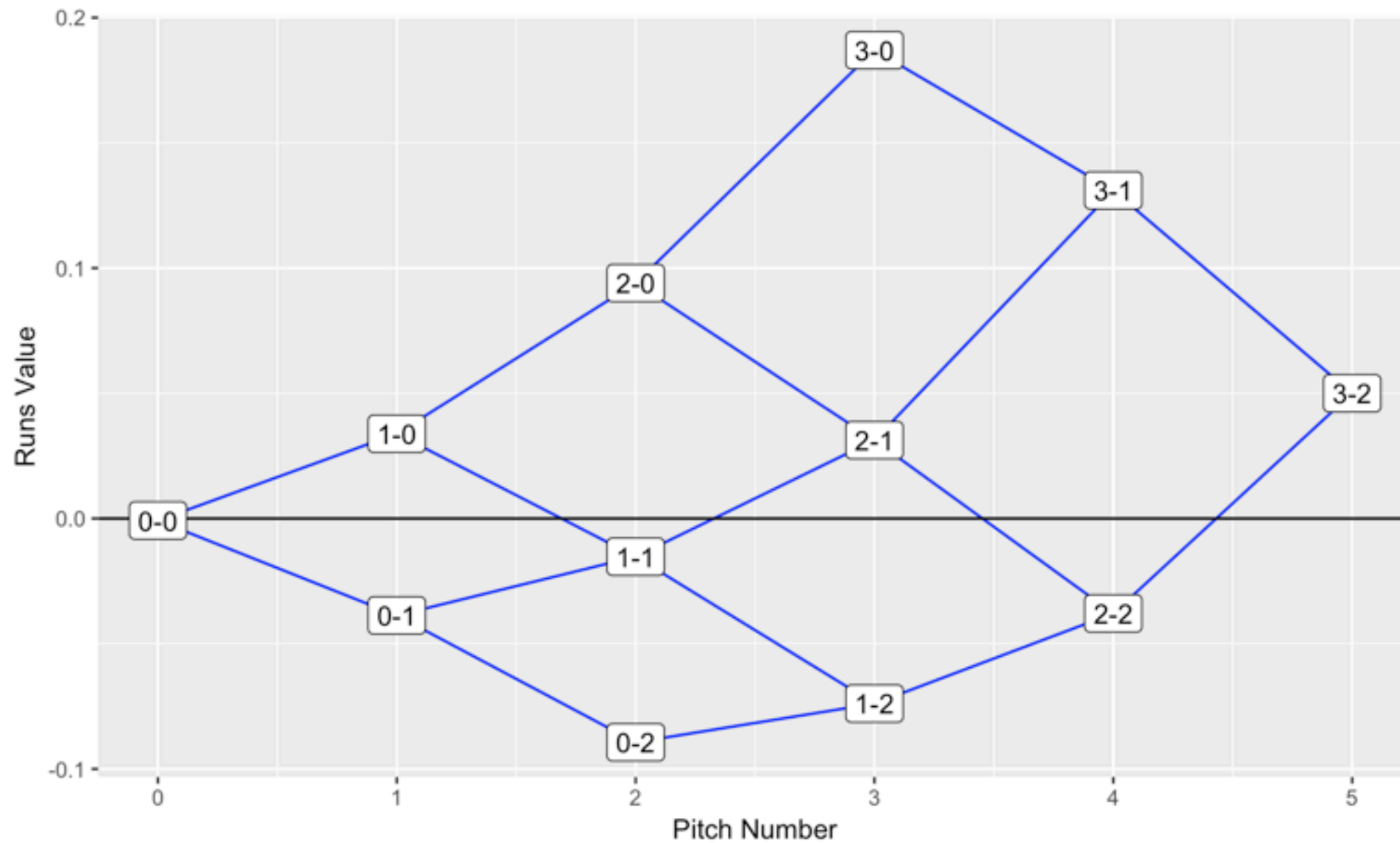
- Mickey Mantle vs Mike Schmidt



기대득점

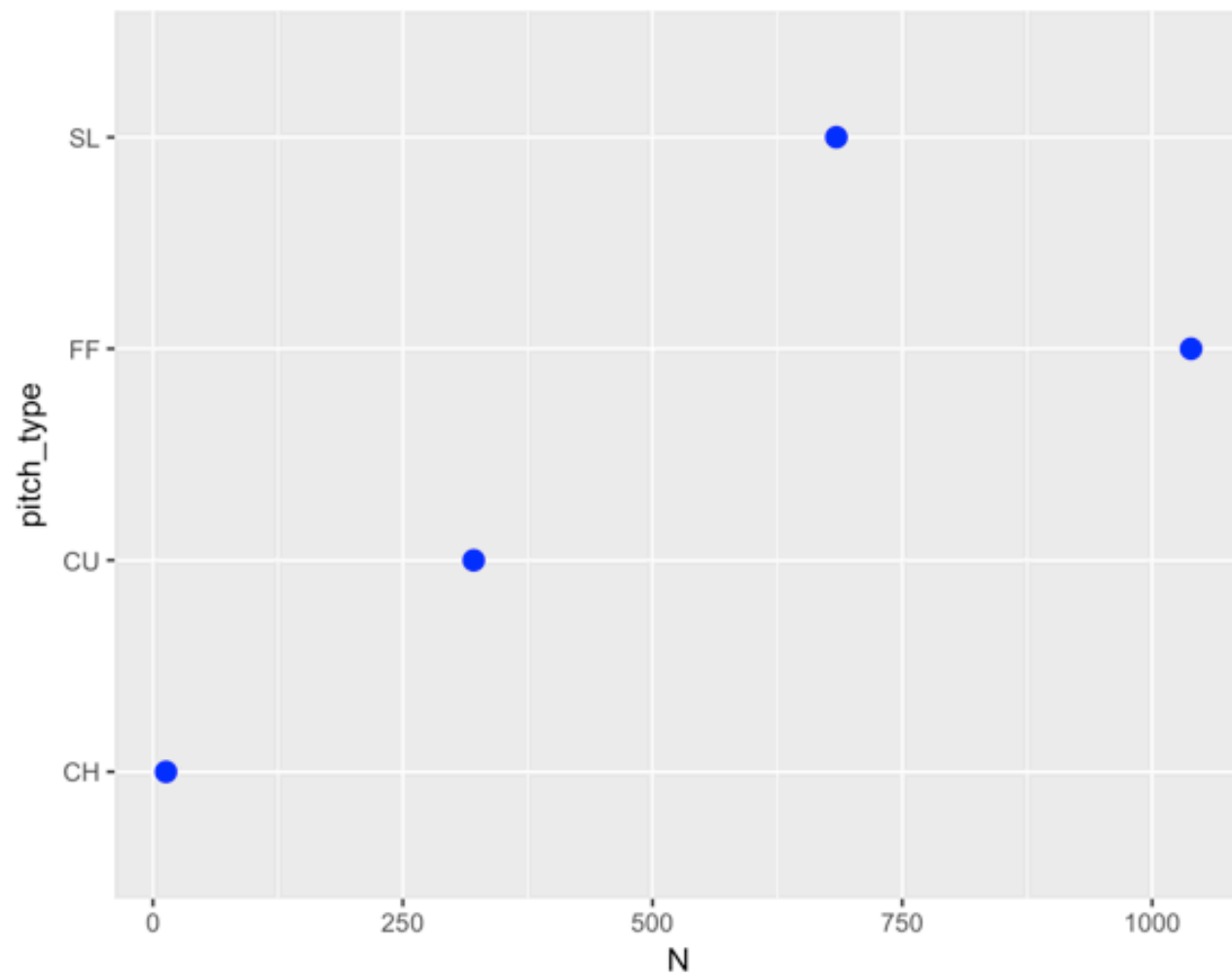


볼카운트별 기대득점

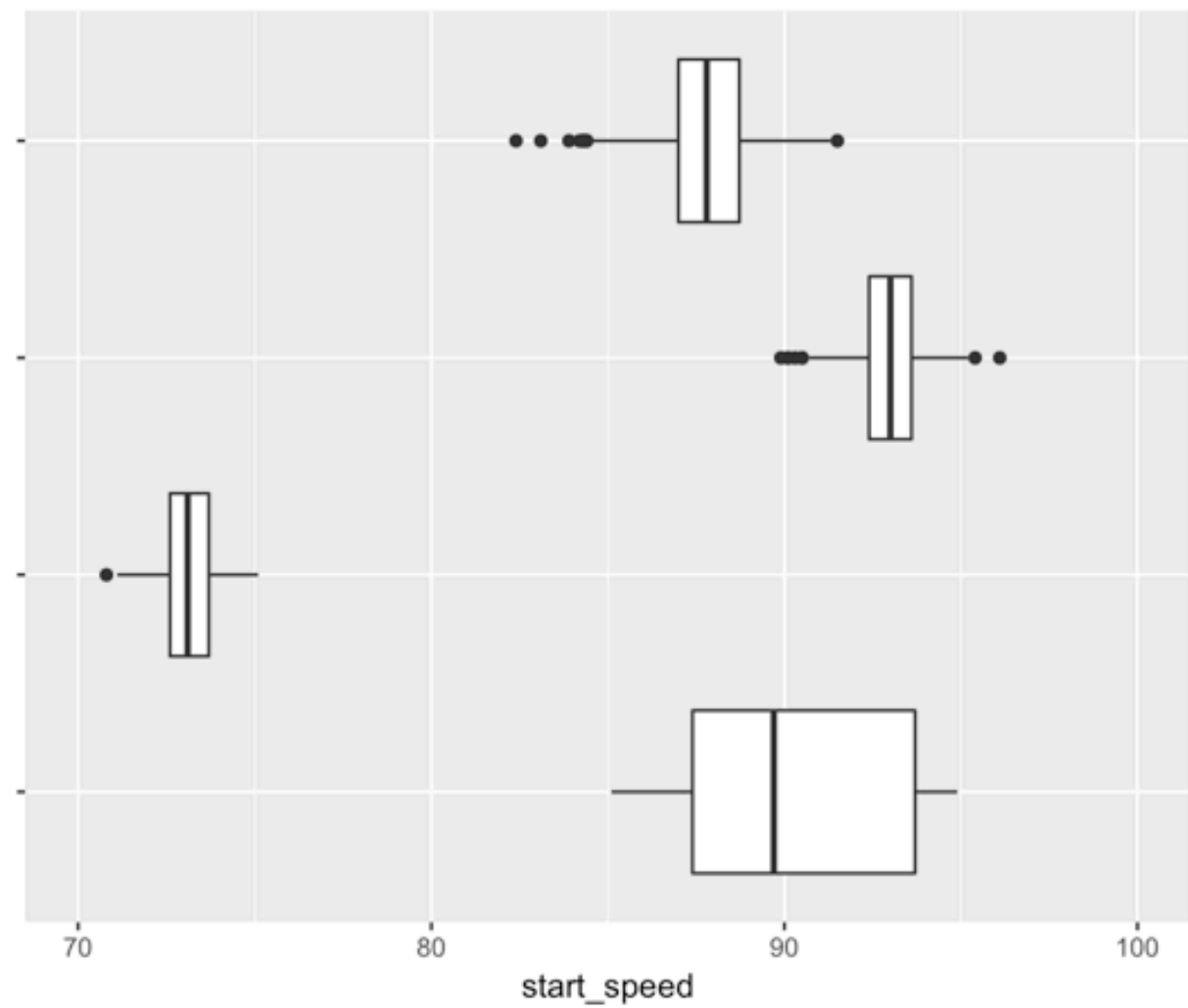


PITCHf/x 자료

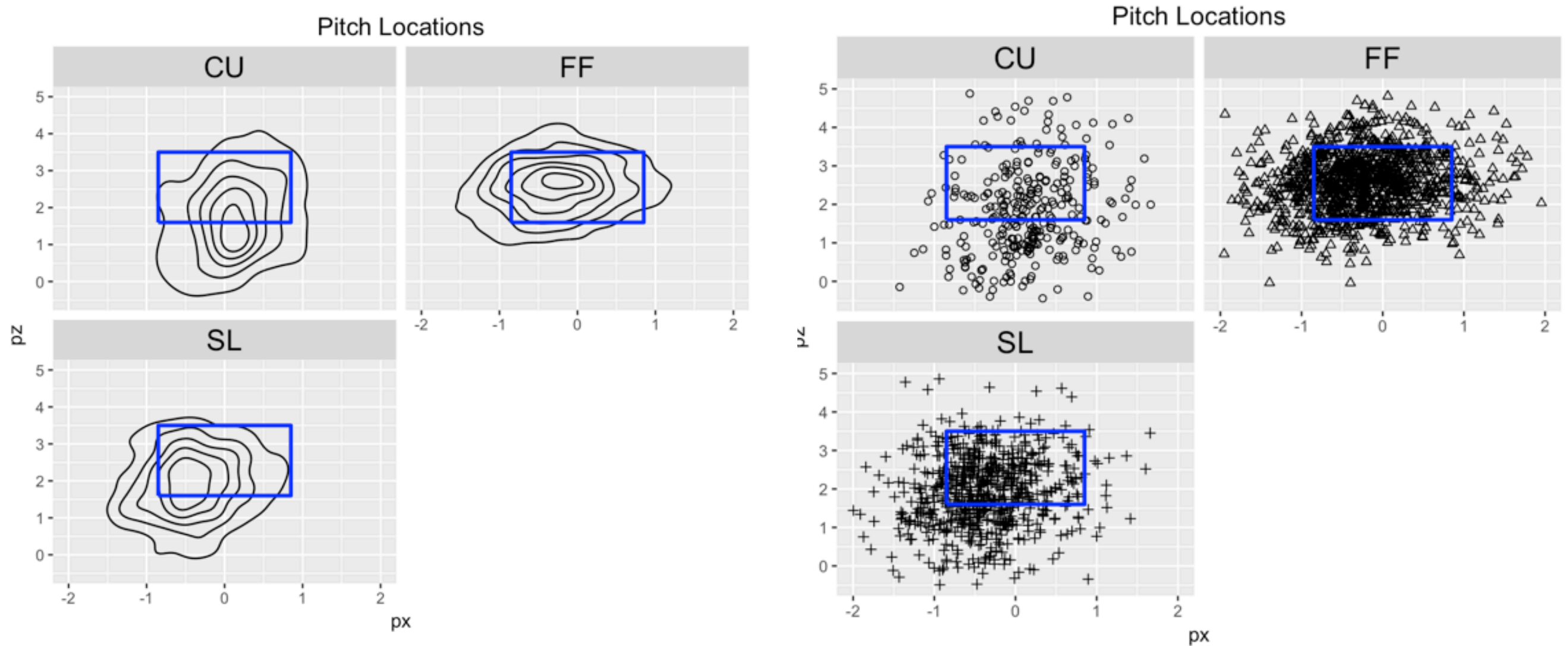
Frequencies of Pitch Type of Clayton Kershaw



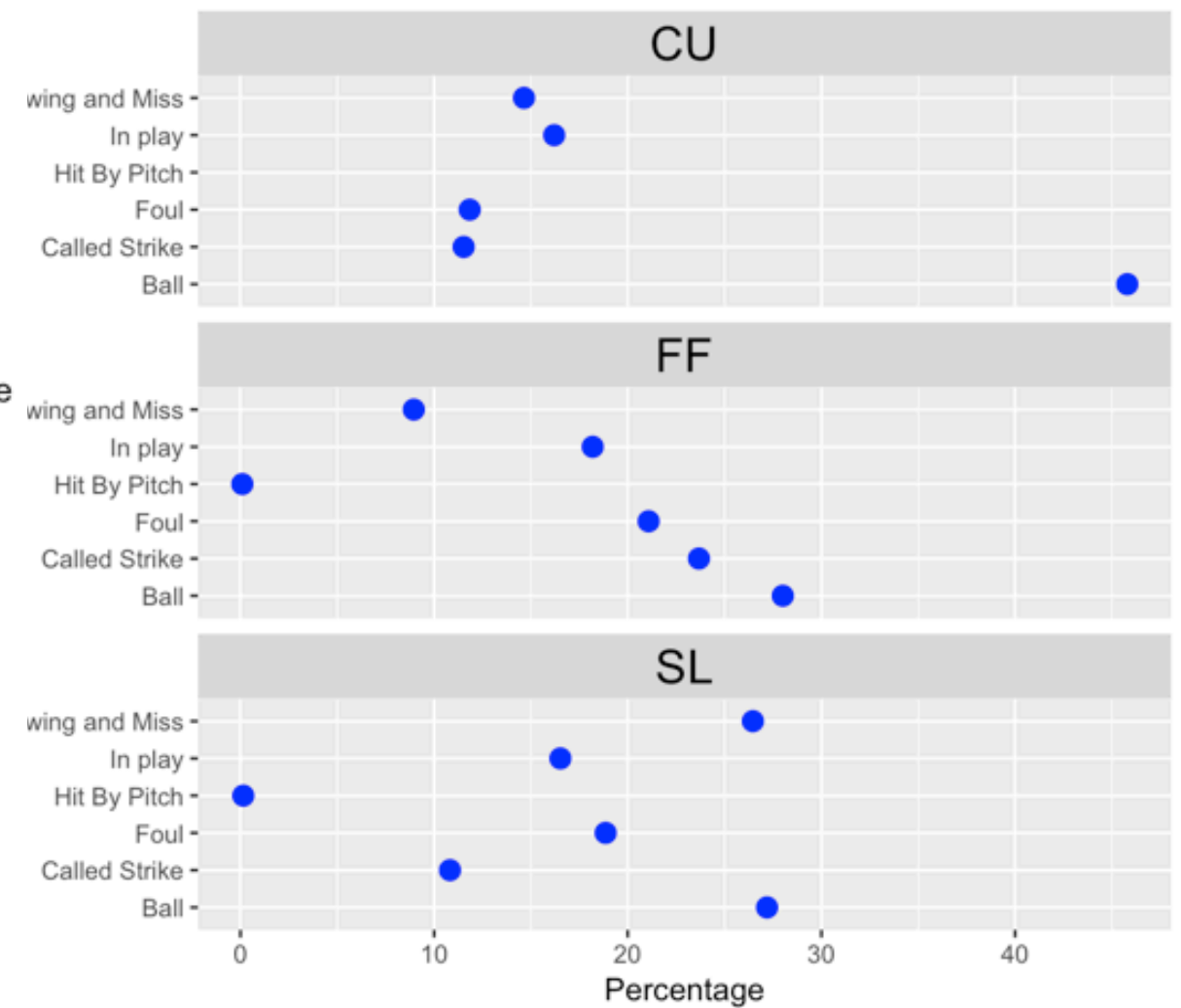
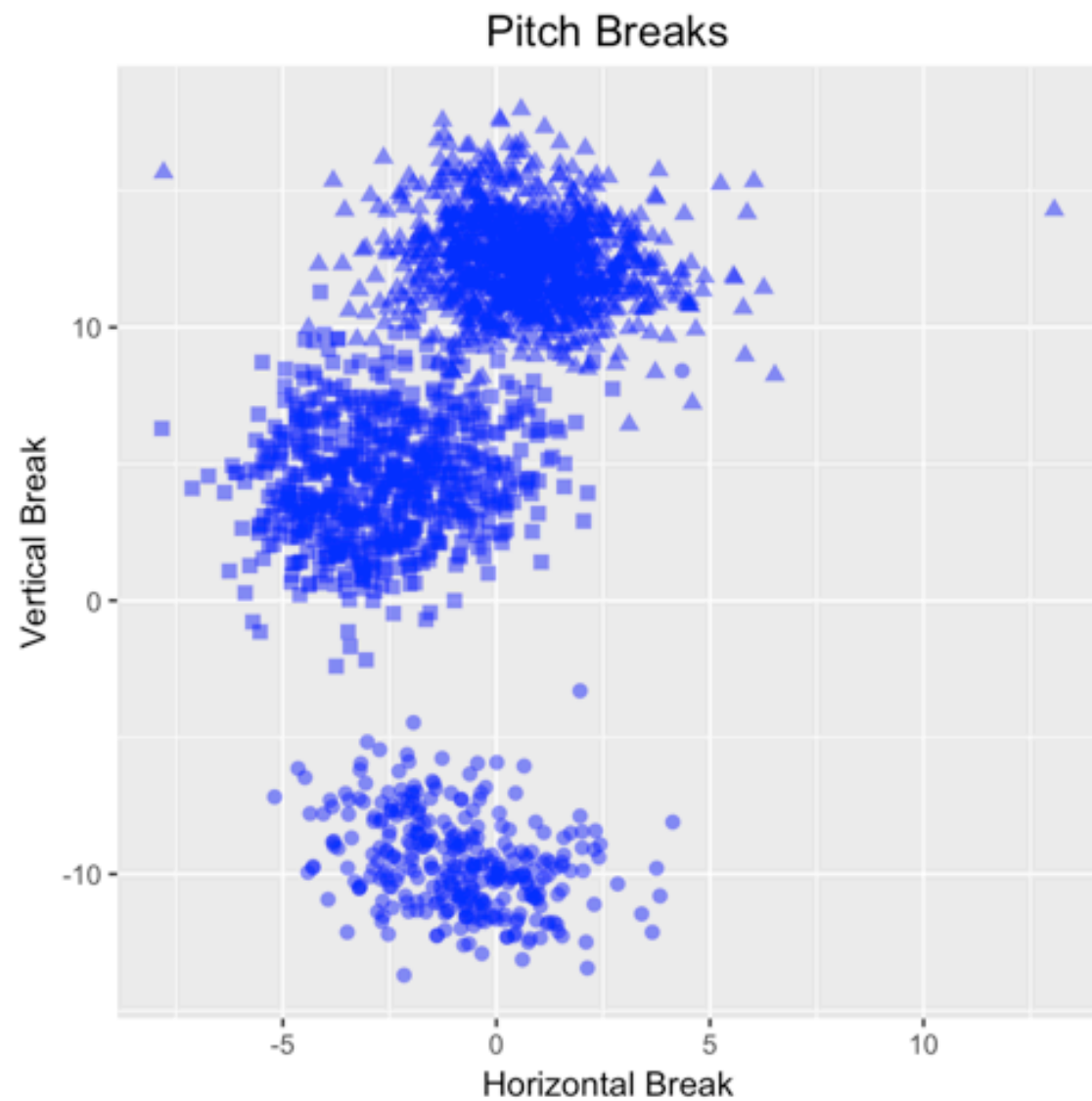
Pitch Speeds



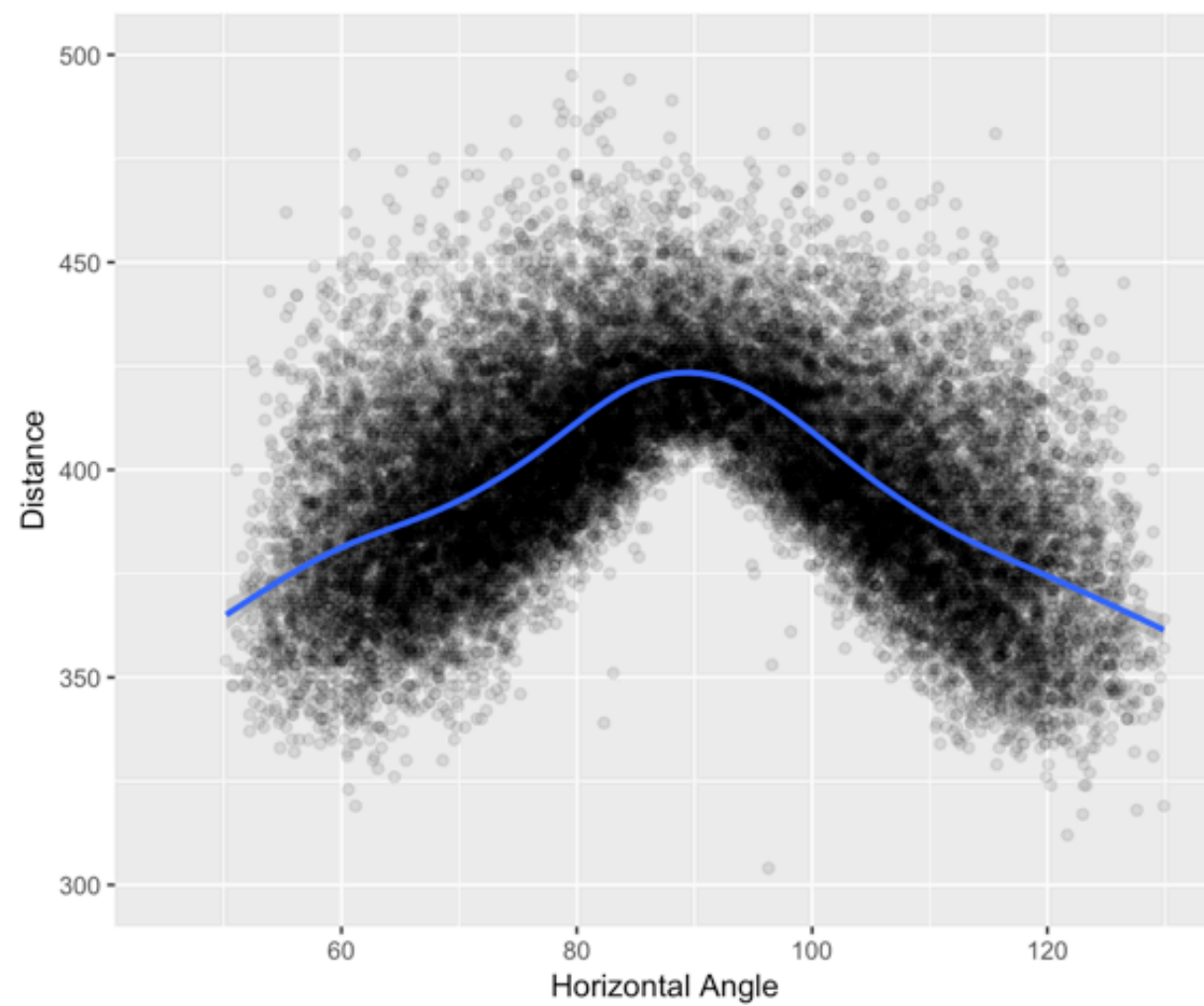
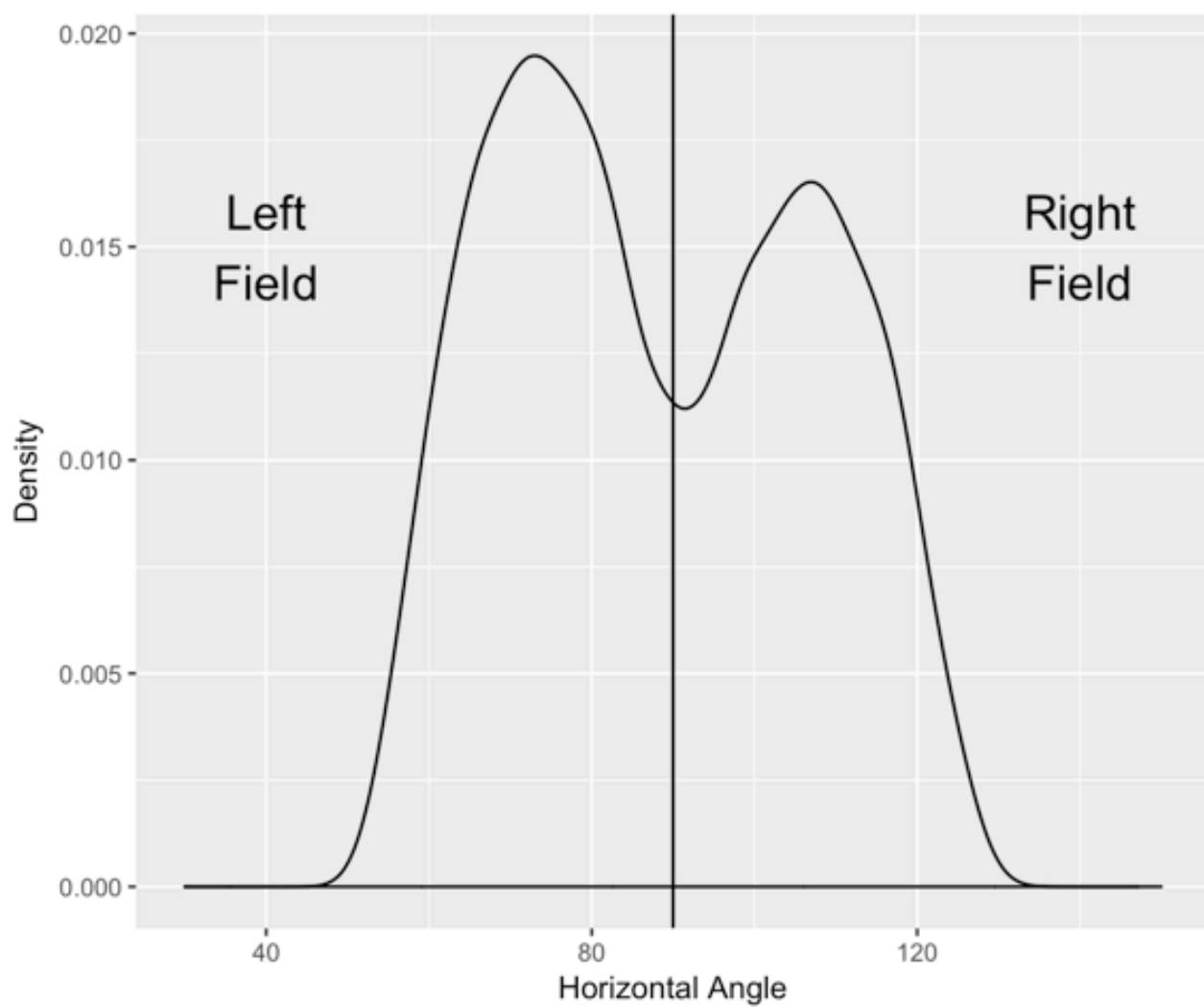
Kershaw's pitch location



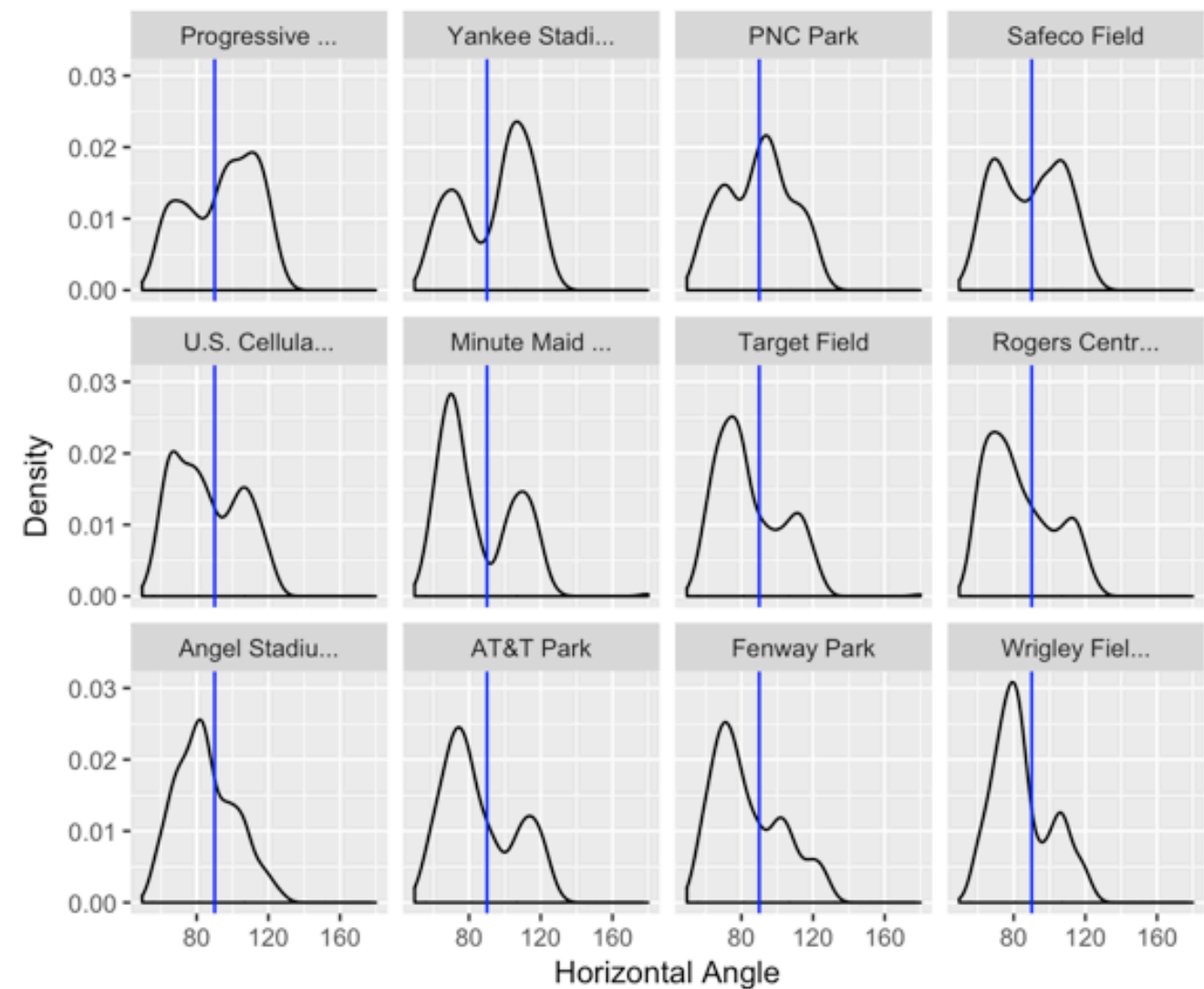
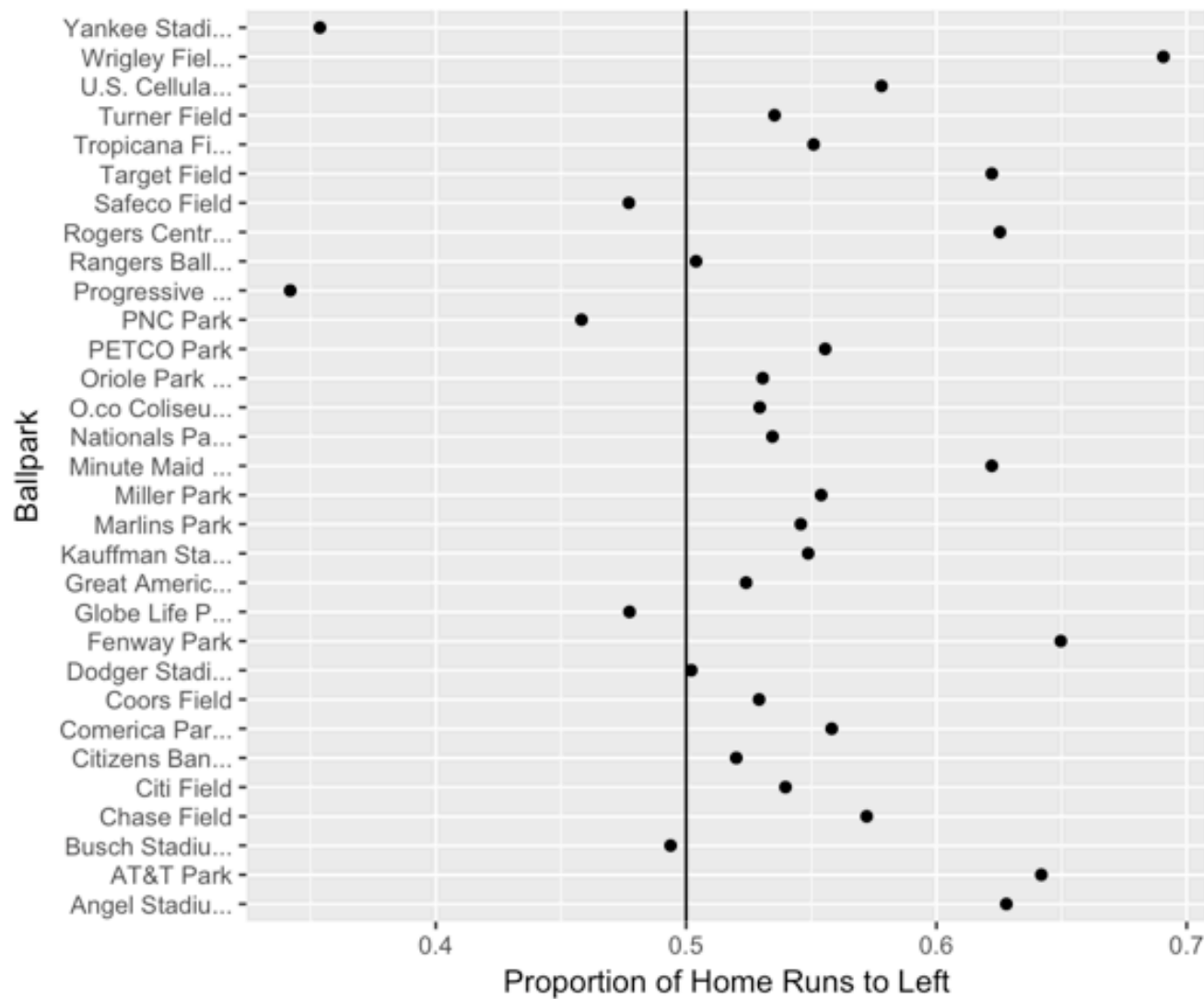
Kershaw's pitch movement



홈런과 타구방향



구장이 홈런에 미치는 효과



하이라이트 편집?

- 108년만에 시카고 컵스가 월드시리즈 우승
- 7차전에서 양팀 합계 총 96번의 타석

MLB · 16. 11. 3.

종료



7

-

8



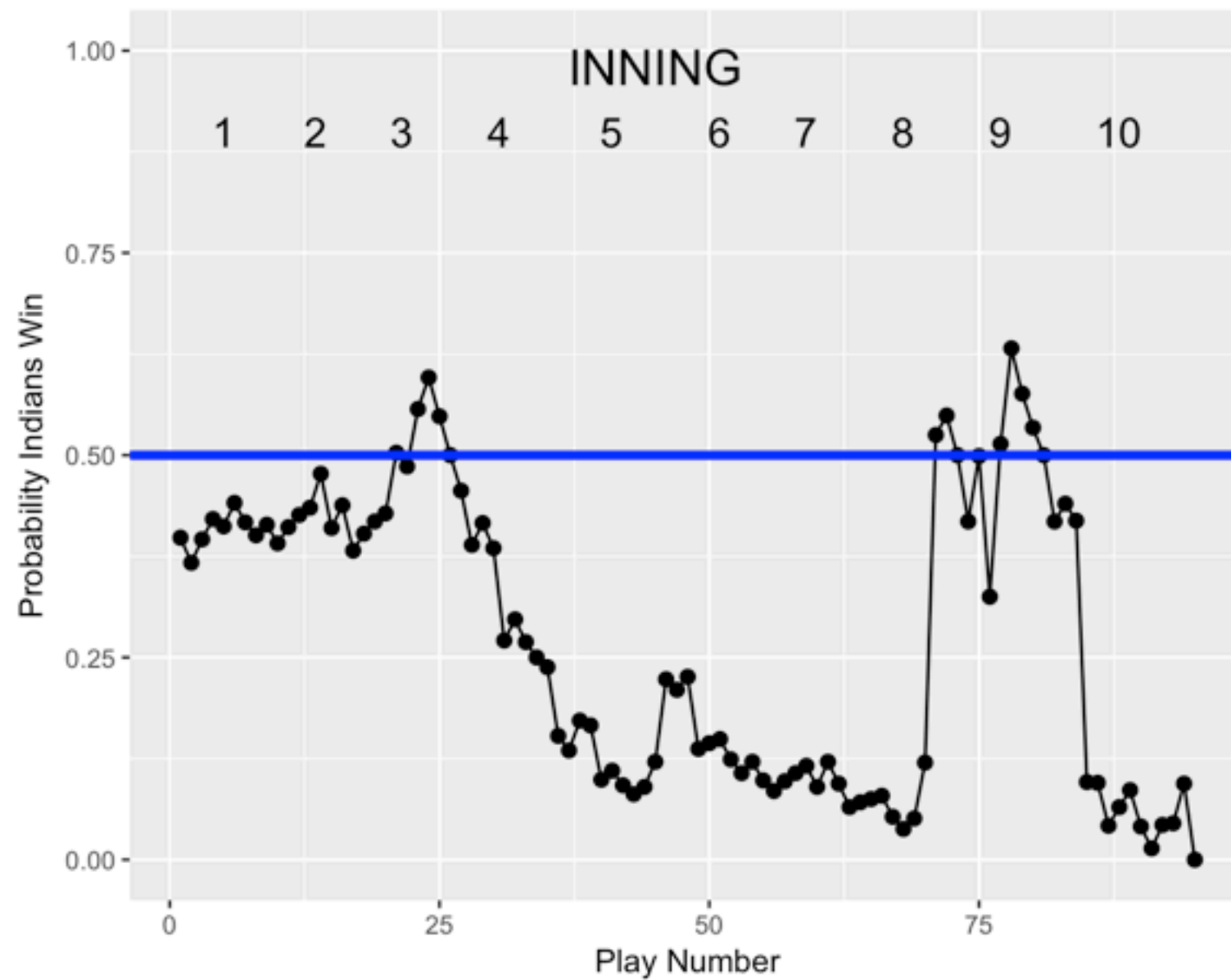
클리블랜드 인디언스
(3승 4패)

시카고 컵스
(4승 3패)

월드 시리즈 · 시카고 C 승리, 4승 3패

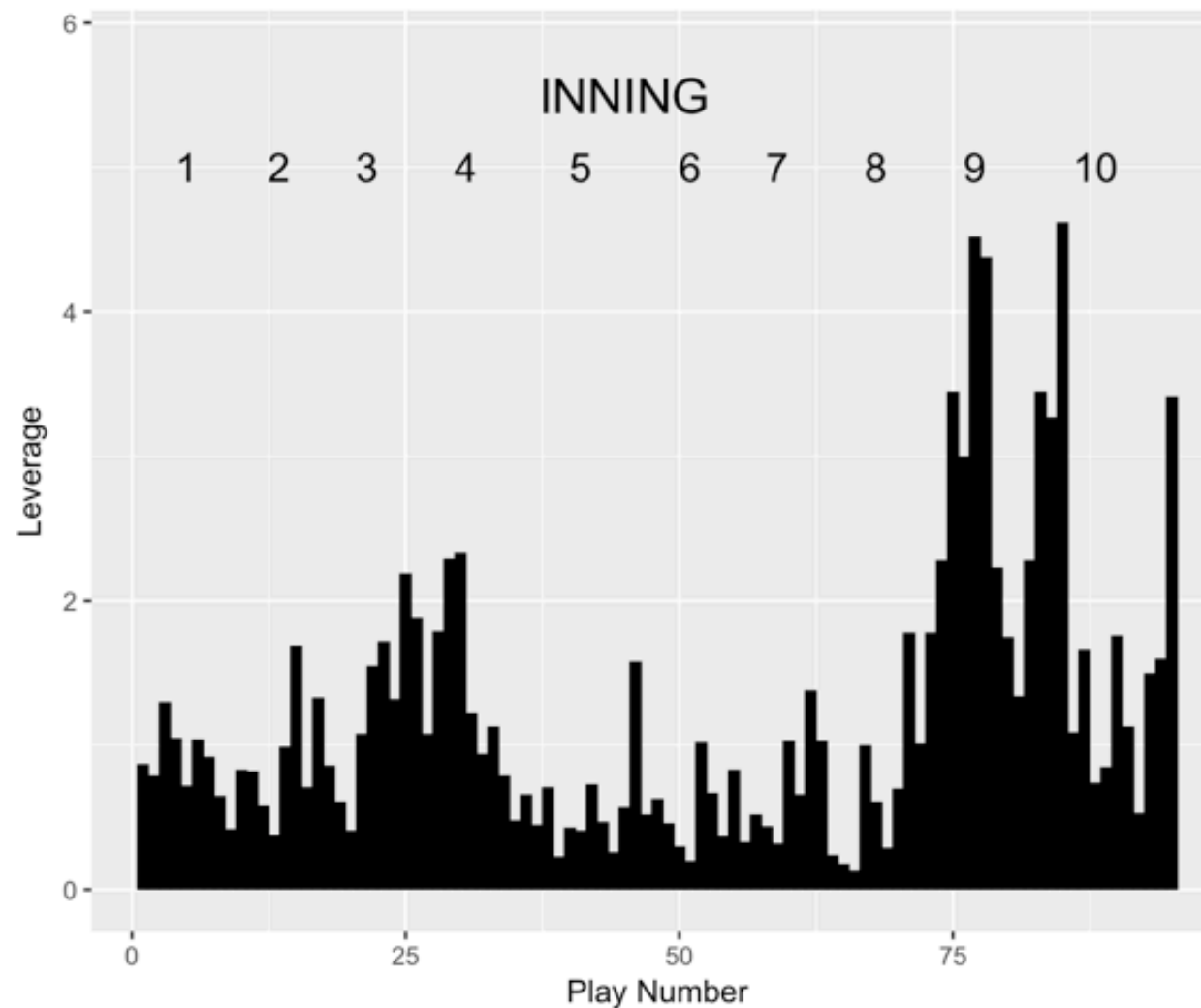
팀	1	2	3	4	5	6	7	8	9	10	R	H	E
시카고 컵스	1	0	0	2	2	1	0	0	0	2	8	13	3
클리블랜드 인디언스	0	0	1	0	2	0	0	3	0	1	7	11	1

홈팀이 이길 확률

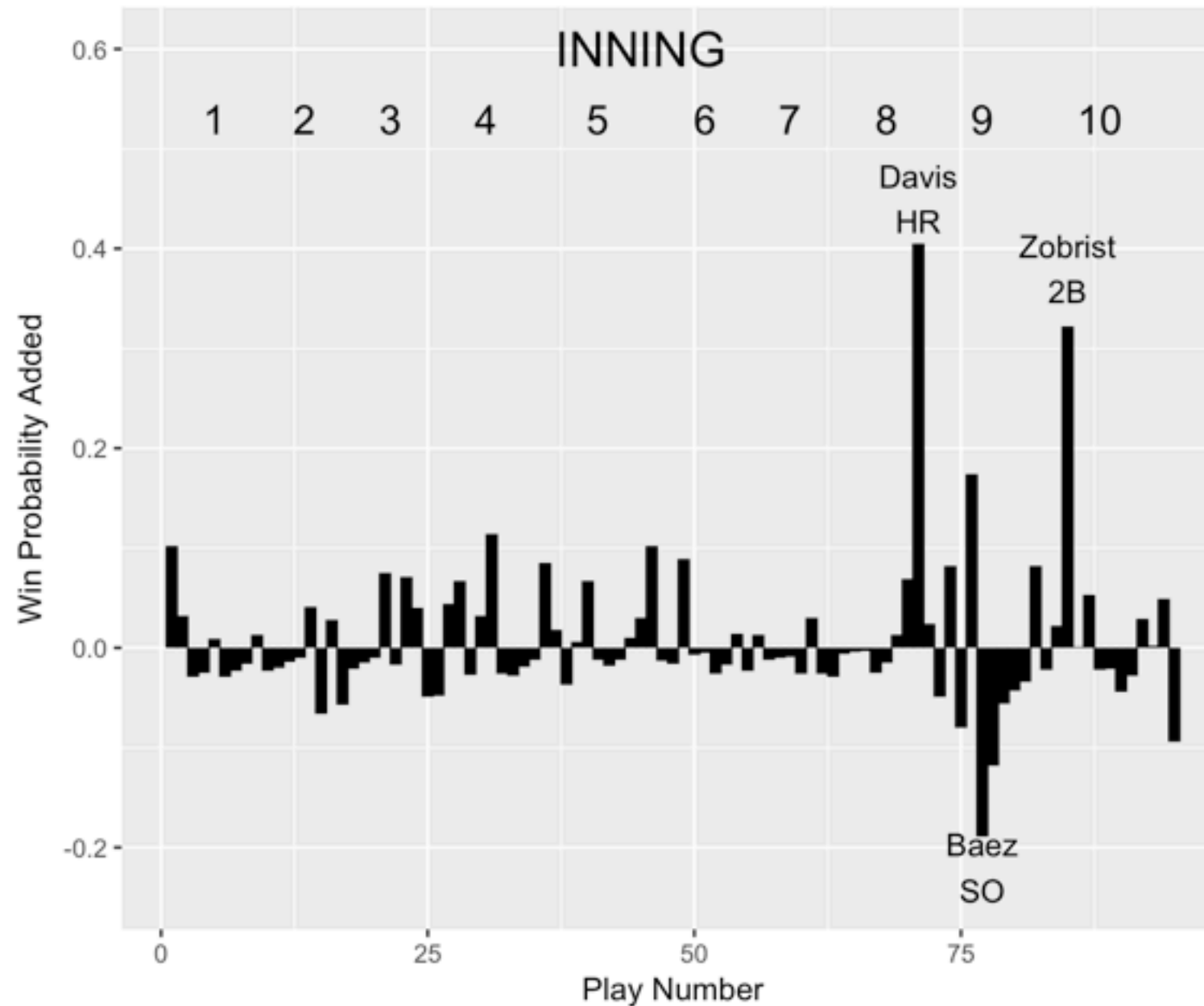


Leverage

- 현재 점수, 이닝, 주자상황,아웃카운트에 따라서 현재 타석의 결과가 홈팀 승율에 영향을 주는 정도



개별 타석별 승률에 미치는 영향



Streaky Hitting

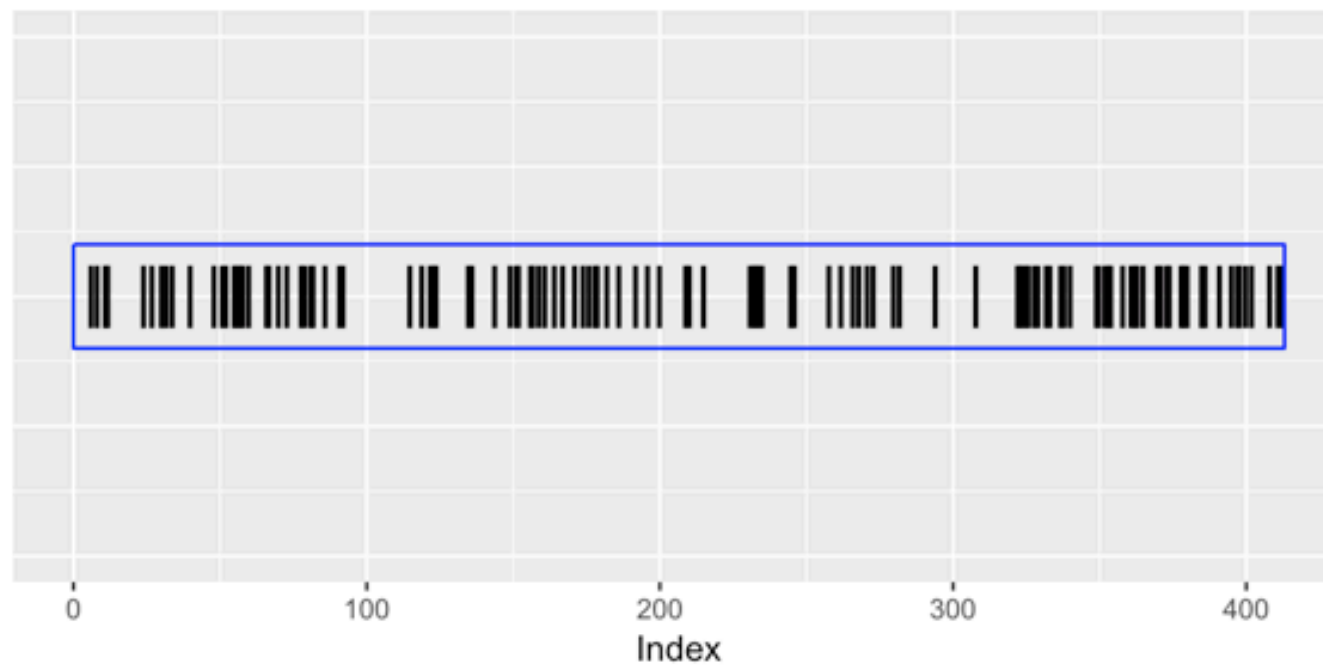
- 2016년 Neil Walker와 Nori Aoki는 타율로만 비교시 매우 비슷한 성적을 냈다
- Walker : $116/412 = 0.282$
- Aoki: $118/417 = 0.283$

Aoki vs Walker

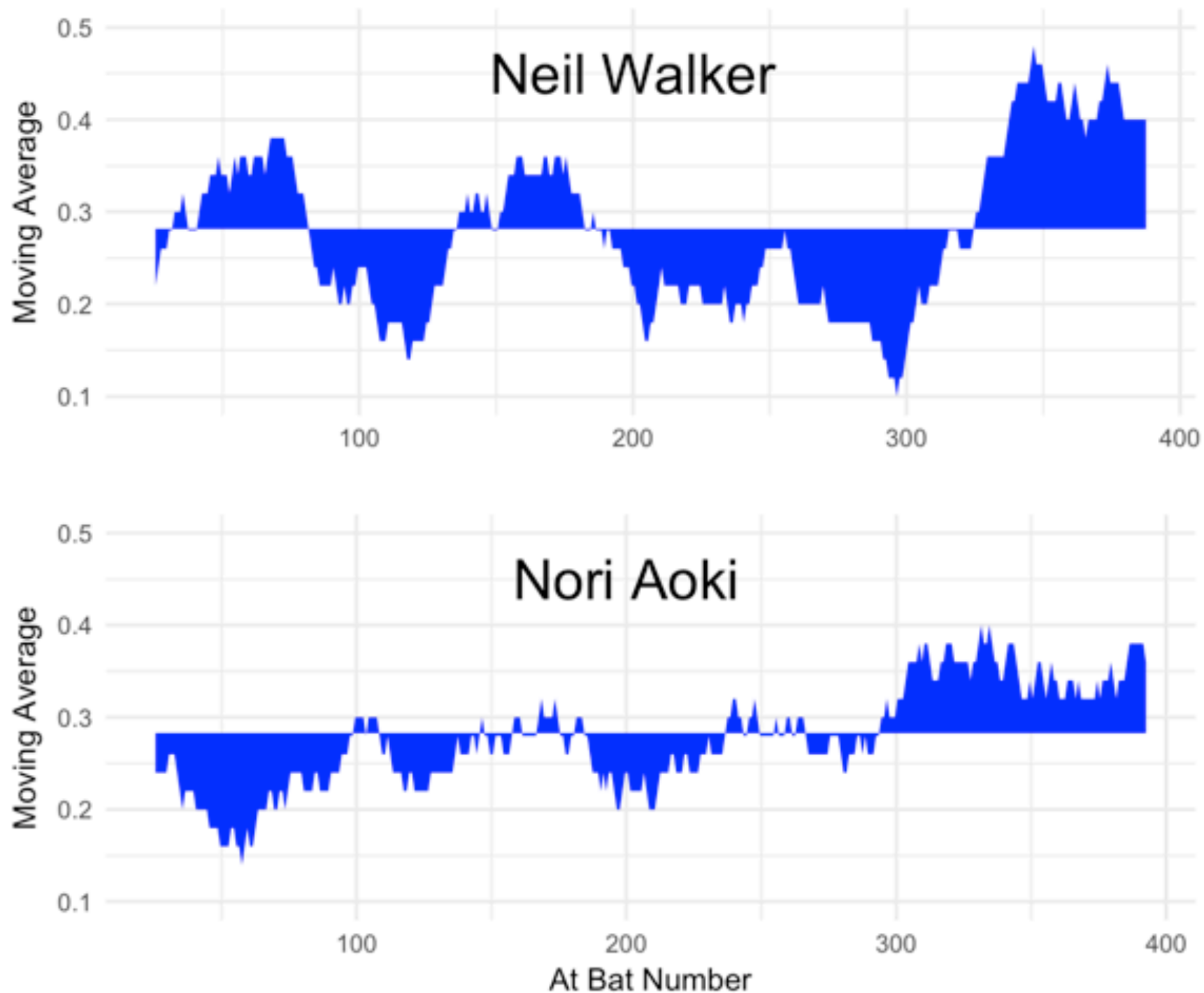
Aoki



Walker

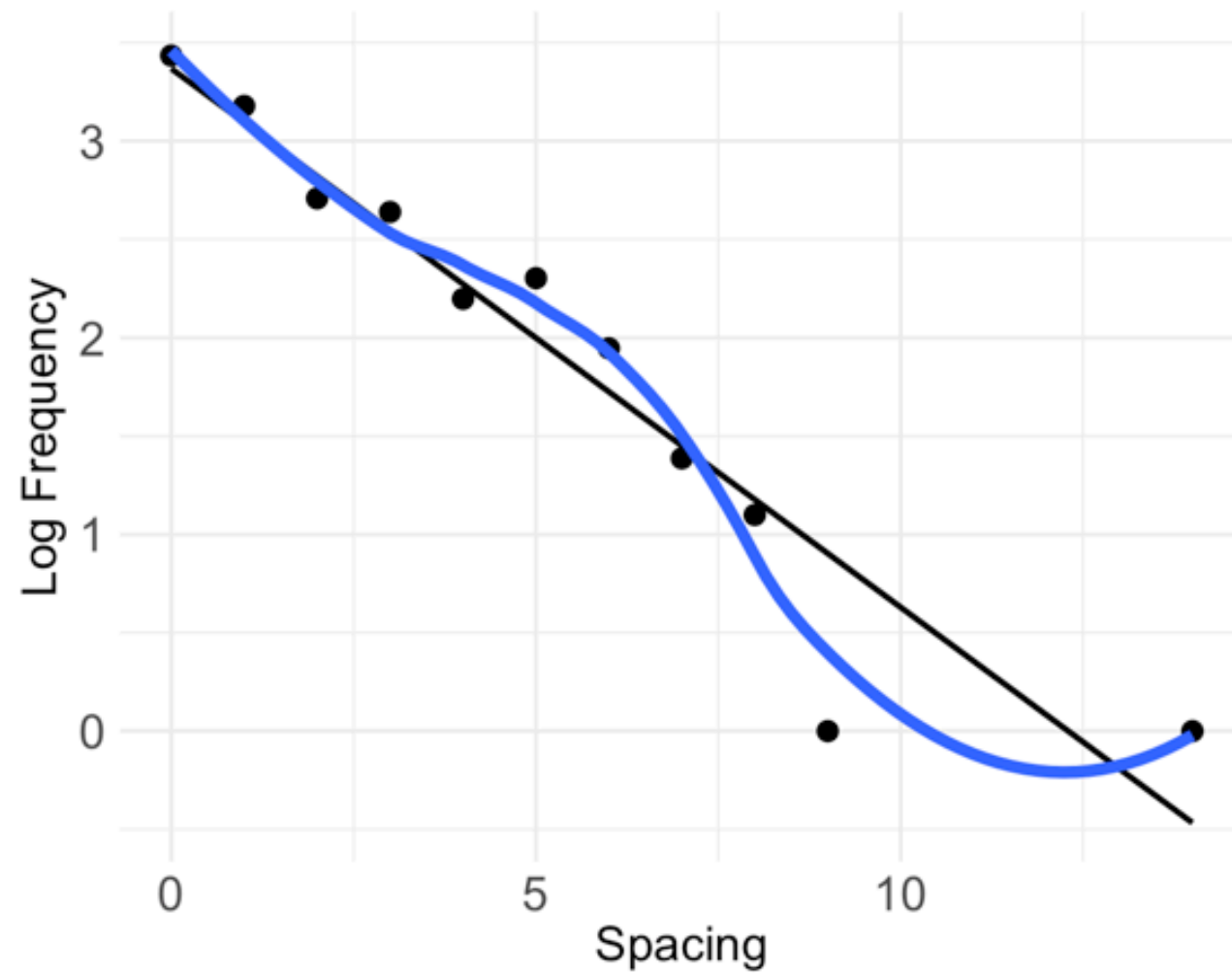


50타석 moving average



Geometric plot

Aoki



Walker

