MLB Umpires – The Study of Imperfection

Jonah Soos

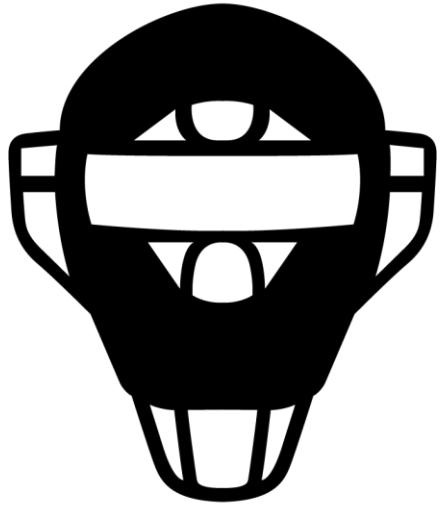
jasoos@syr.edu | linkedin.com/in/jonahsoos24

Goals of the Project

- Analyze the variability of precision and accuracy during decision making
    - Utilize MLB Umpires as a case study
- Look for trends and factors that may cause variability
- View the magnitude of the effects.
- Evaluate possible solutions

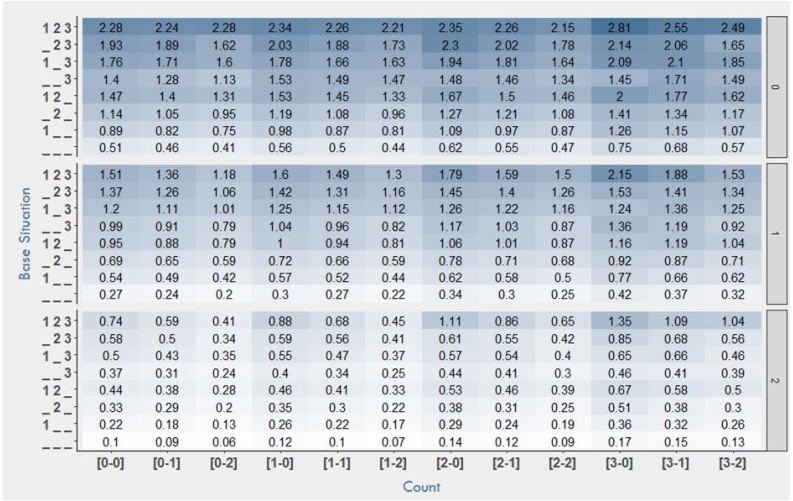
Methodology

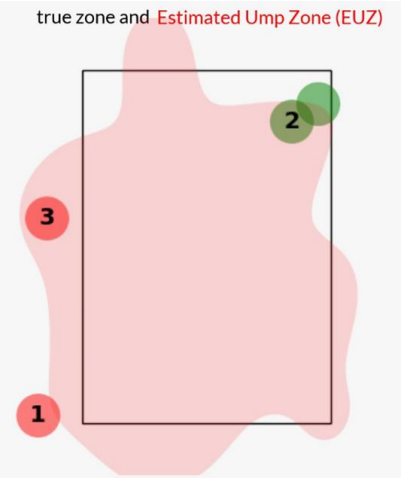
- Collect and cumulate game, umpire, and team data from Umpire Scorecards
- Conduct an analysis of umpire performance using experience as the dependant variable
- Create a Overturn Simulation to see the possible seasonwide affects



Important Terms

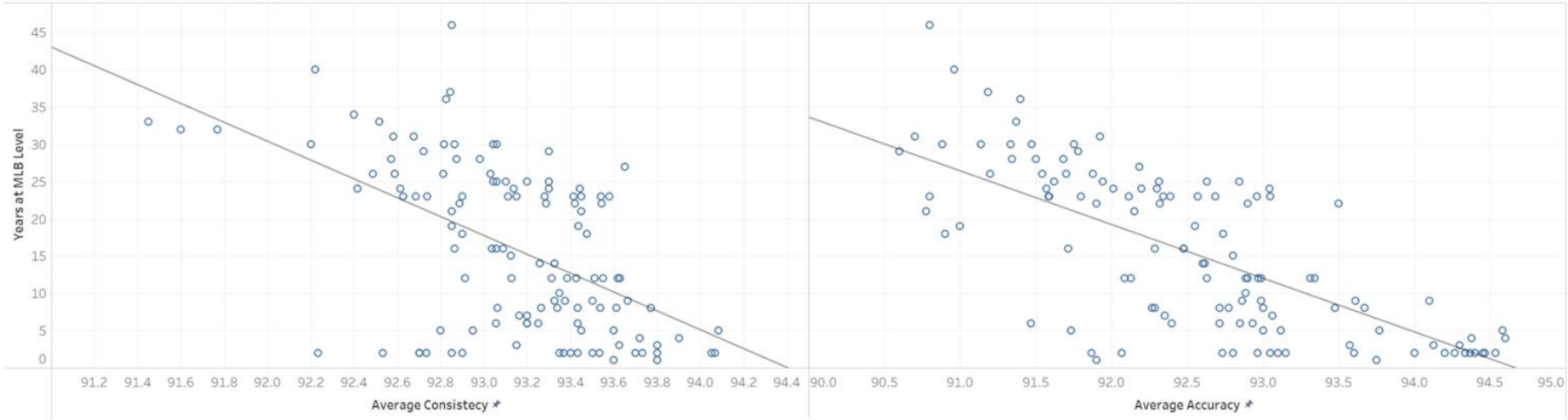
- Accuracy – How accurate an umpire is within the MLB Standard Strikezone
- Consistency – How accurate an umpire is within his own strikezone
- Favor – The cumulative difference in run expectancy values for inaccurate pitches.



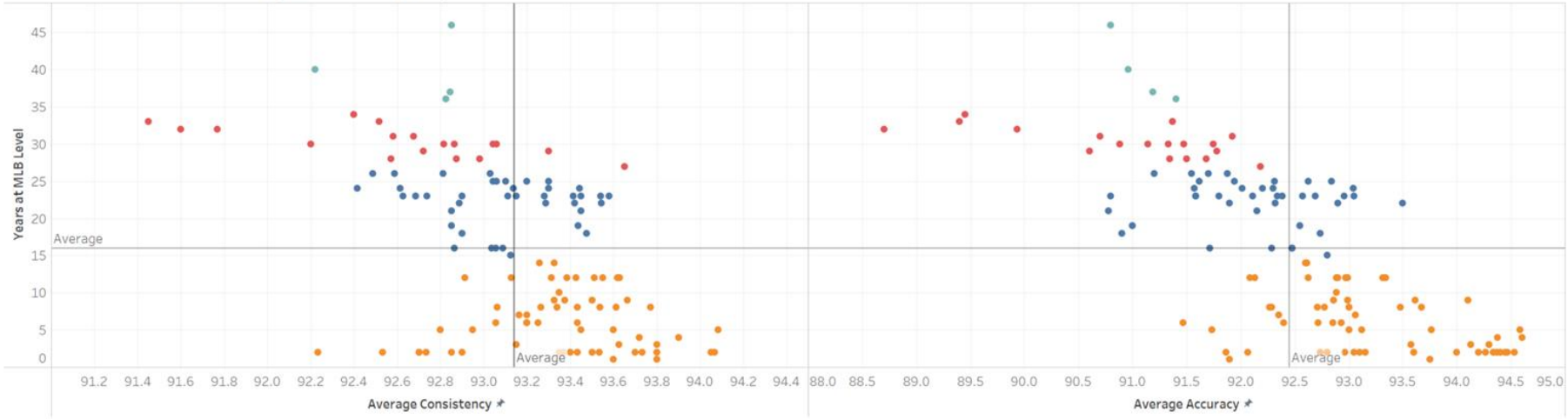


Umpire Analysis

Average Consistecy vs Experience

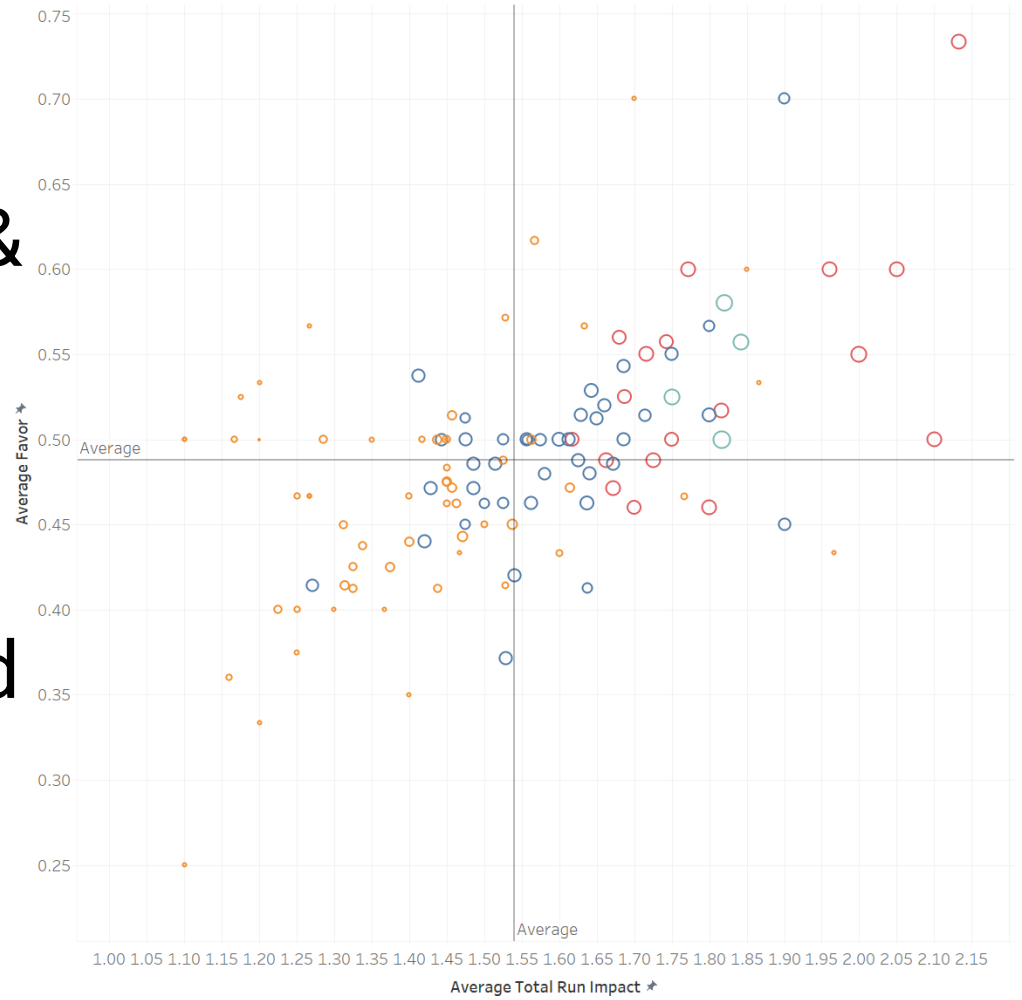


Consistency vs Experience (Cluster Analysis)



Conclusions

- Significant Correlation between Years Experience & Umpire Impact
- Strong case to say experience/age can be used to explain differences in accuracy between umpires.



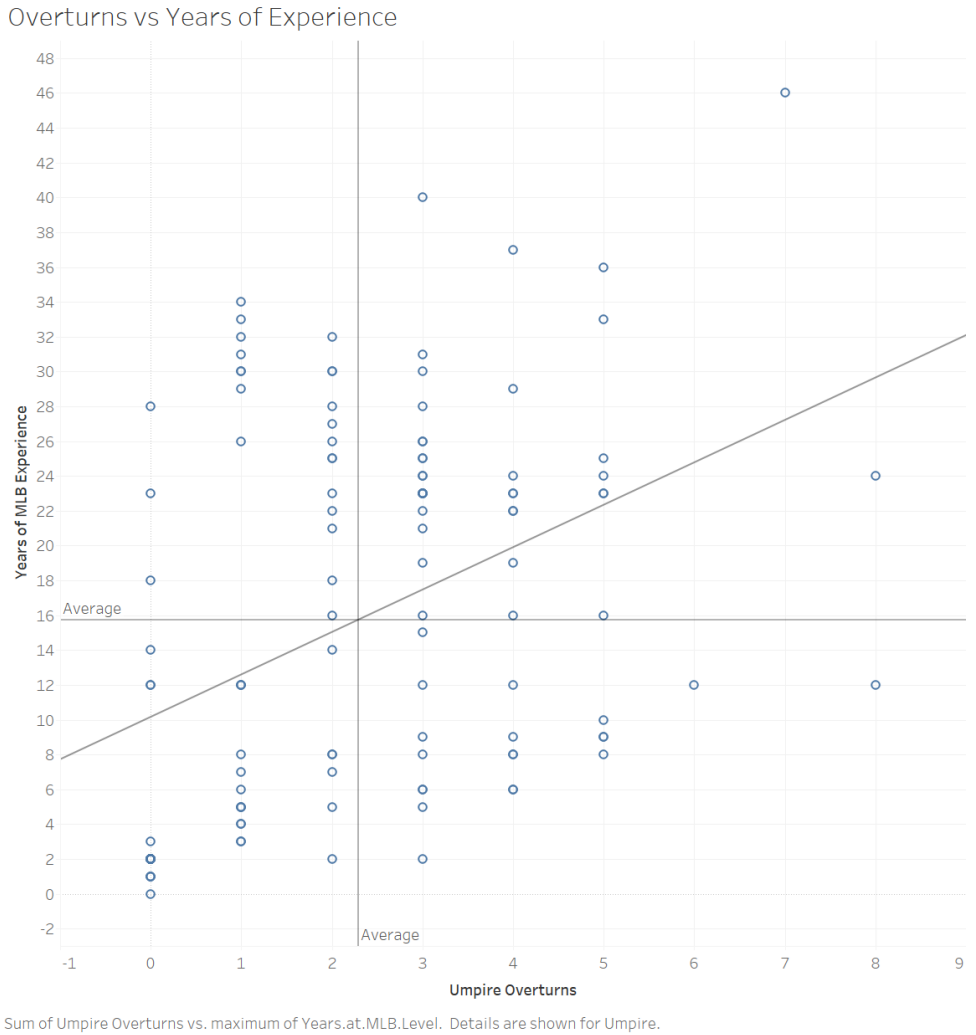
Umpire Overturns

- What is it: When the favor of a team is greater than its winning or losing run differential.
- 284 occurrences ~2% of games since 2015
- 171 vs 113 for the Home and Away teams
- Significant decrease in overturns per year.

Year	Overturns
2015	57
2016	49
2017	48
2018	31
2019	35
2020	20
2021	22
2022	22

Umpire Impact

- Positive Correlation between experience and Overturns (R^2 = .16)
- Fits similar theme



Season Wide Impact

- Separated Overturns into wins gained and lost per team
- Calculated percentage of overturned outcome using pythagorean expected win percentage taking into account favor

Team	Wins Gained	Wins Lost	Team PlusMinus
ARI	11	10	+1
ATL	5	13	-8
BAL	8	5	+3
BOS	9	20	-11
CHC	12	11	+1
CIN	20	5	+15
CLE	4	12	-8
COL	7	6	+1
CWS	8	8	0
DET	8	3	+5
HOU	14	9	+5
KC	10	4	+6
LAA	6	11	-5
LAD	4	10	-6
MIA	9	10	-1
MIL	8	10	-2
MIN	7	10	-3
NYM	9	11	-2
NYY	16	10	+6
OAK	12	7	+5
PHI	9	2	+7
PIT	8	4	+4
SD	9	12	-3
SEA	9	18	-9
SF	12	10	+2
STL	11	11	0
TB	6	9	-3
TEX	8	8	0
TOR	10	11	-1
WSH	15	14	+1

Case Study 1

2019 NL Playoff Standings				
Seed		Team	Record	Wild Card
1	LA	Los Angeles Dodgers	106-56	--
2	A	Atlanta Braves	97-65	--
3	STL	St. Louis Cardinals	91-71	--
4	W	Washington Nationals	93-69	+4
5	MIL	Milwaukee Brewers	89-73	0
E	NYM	New York Mets	86-76	-3
E	AZ	Arizona Diamondbacks	85-77	-4

2019 NL Adjusted Playoff Standings				
Seed		Team	Record	Wild Card
1	LA	Los Angeles Dodgers	106-56	--
2	W	Washington Nationals	95-67	--
3	A	Atlanta Braves	94-68	+6
4	STL	St. Louis Cardinals	91-71	--
5	MIL	Milwaukee Brewers	88-74	0
E	AZ	Arizona Diamondbacks	86-76	-2
E	NYM	New York Mets	84-78	-4
		Probability of Scenario:	7.90%	

Outcome	Probability
WSH BEAT ATL	32.47%
WSH TIE ATL	23.35%
ATL BEAT WSH	47.39%

Case Study 2

2018 NL Playoff Standings			
Seed	Team	Record	Wild Card
T-1	Milwaukee Brewers	95-67	--
T-1	Chicago Cubs	95-67	+4
T-2	Los Angeles Dodgers	91-71	--
T-2	Colorado Rockies	91-71	0
3	Atlanta Braves	90-72	--
E	St. Louis Cardinals	88-74	-3
E	Pittsburgh Pirates	82-79	-9

2018 NL Adjusted Playoff Standings			
Seed	Team	Record	Wild Card
1	Milwaukee Brewers	95-67	--
2	Atlanta Braves	91-71	--
3	Colorado Rockies	90-73	--
4	Chicago Cubs	94-68	+6
5	St. Louis Cardinals	89-74	0
E	Los Angeles Dodgers	88-74	0
E	Pittsburgh Pirates	81-80	-7

Outcome	Probability
CHC BEAT MIL	67.52%
CHC TIE MIL	16.36%
MIL BEAT CHC	10.44%
COL BEAT LAD	45.61%
COL TIE LAD	32.80%
LAD BEAT COL	21.58%

Outcome	Probability
STL ADV, LAD ELIM	14.25%
STL & LAD TIE	25.43%
STL & COL TIE	12.52%
STL ELIM	30.30%

Case Study 3

2016 NL Playoff Standings			
Seed	Team	Record	Wild Card
1	Chicago Cubs	103-59	--
2	Washington Nationals	95-67	--
3	Los Angeles Dodgers	91-71	--
T-4	New York Mets	87-75	0
T-4	San Francisco Giants	87-75	0
E	St. Louis Cardinals	86-76	-1
E	Miami Marlins	79-82	-8

2016 NL Adjusted Playoff Standings			
Seed	Team	Record	Wild Card
1	Chicago Cubs	101-61	--
2	Washington Nationals	95-67	--
3	Los Angeles Dodgers	91-71	--
T-4	New York Mets	86-76	0
T-4	St. Louis Cardinals	86-76	0
T-4	San Francisco Giants	86-76	0
E	Miami Marlins	80-81	-6
	Probability of Scenario:	17.74%	

Outcome	Probability
NYM & SF ADV	20.54%
SF & STL ADV	12.53%
NYM & STL ADV	16.18%
3 WAY TIE	17.74%
STL & SF TIE, NYM ADV	6.76%
STL & NYM TIE, SF ADV	14.60%
SF & NYM TIE, STL ADV	10.31%

Concluding Thoughts

- Umpire variability is legitimate, and correlates strongly with years of experience.
- No bias was found towards or against any team
- Variability can have a real impact on a season.
- What will the effect of AAA Robot Umpires be?

Thank you!

Questions?

Contact info: jasoos@syr.edu | linkedin.com/in/jonahsoos24