Comparing Baseball Player Performance Metrics: NCAA Division 1 vs. Frontier League

Zac Ambrose, Robert Cowan, Nathaniel Ascher, Youssef Gehad

Problem statement

- MLB teams have many different leagues to scout player talent from

- How can professional teams be sure that the players they are acquiring are the best available, regardless of which league they play in?

Research question

How does player performance differ across leagues?

- Could help MLB teams properly select talent based on performance in lower leagues

- We are looking to answer this question with 2023 data from NCAA Division 1 teams and the Frontier League, an independent summer league

- Access provided for data for the 2023 NCAA season by Rice Baseball analytics and 2023 Frontier League season by Nathaniel Ascher

How does player performance differ across leagues?

Text NCAA Frontier

- Typically, players in NCAA are regarded as better than those in the Frontier League
- If one league is “better” on average, we’d expect to see improved player performance metrics
- We will focus on pitching metrics, as they are more comparable across leagues

Random forest

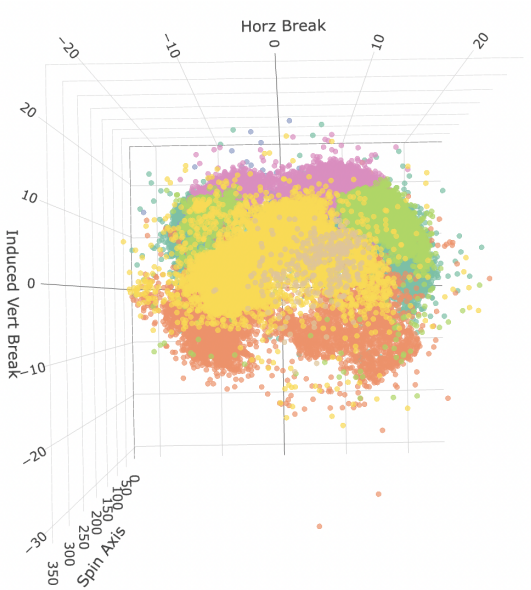
- While the NCAA data had an auto pitch type, the Frontier League data did not
- We used a random forest model to predict the pitch type for the Frontier League data using the distribution of NCAA pitch types as the training data
- Features for the model included induced vertical break, horizontal break, release speed, spin rate, spin axis, and pitcher's handedness to predict pitch type
- Resulted in 96.6% accuracy on the test data

Pitch metric definitions

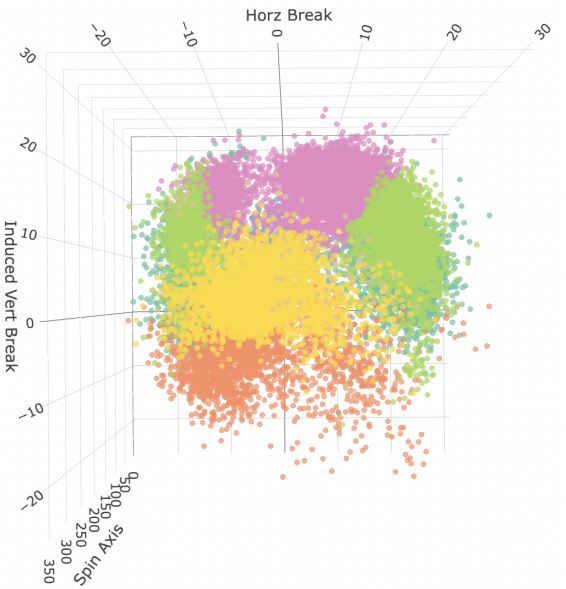
Horizontal Break (HB)

Induced Vertical Break (IVB)

- Changeup
- Curveball
- Cutter
- Fastball
- Sinker
- Slider
- Splitter

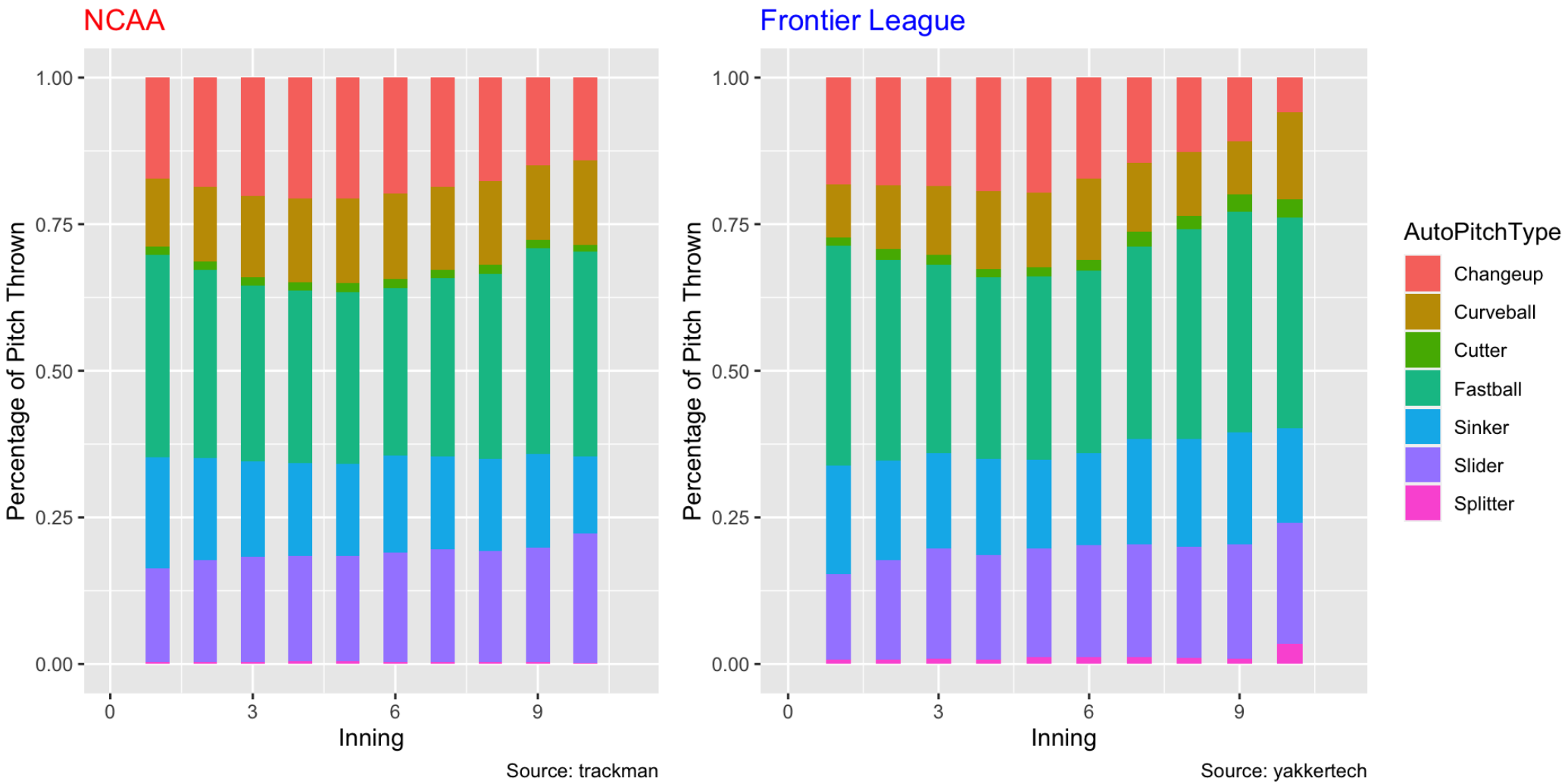


Frontier

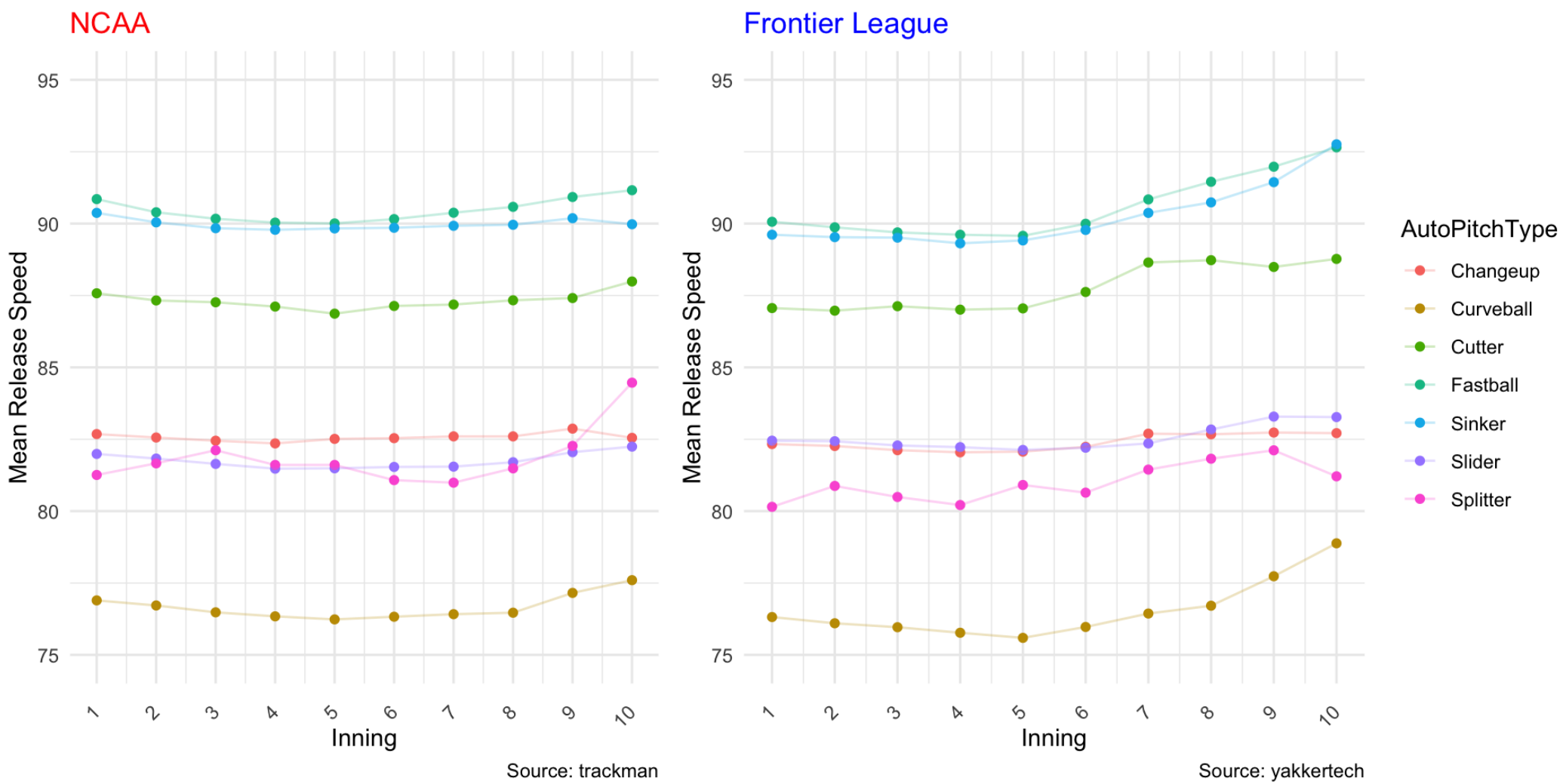


NCAA

Pitch type by inning

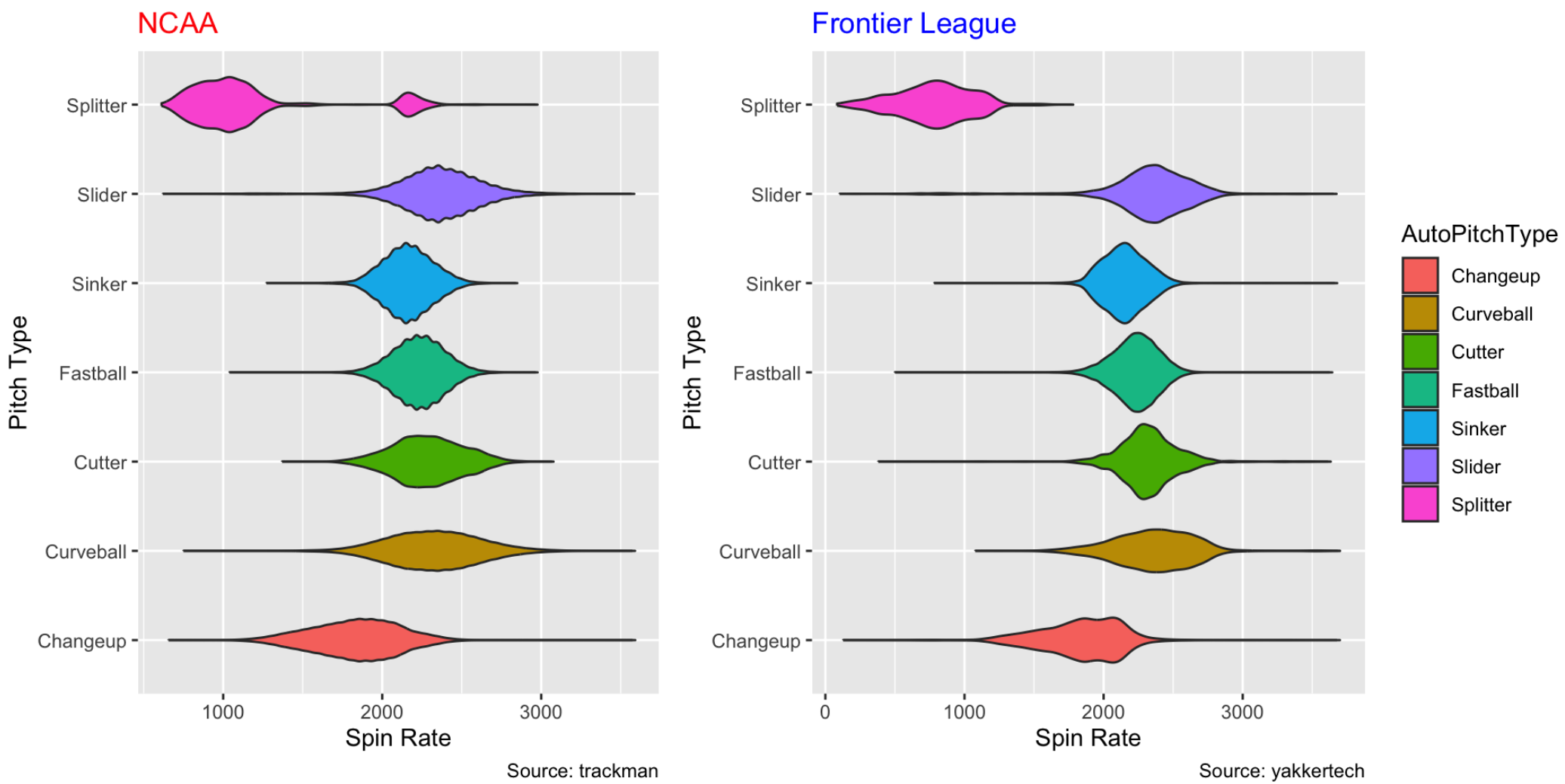


Release speed by inning



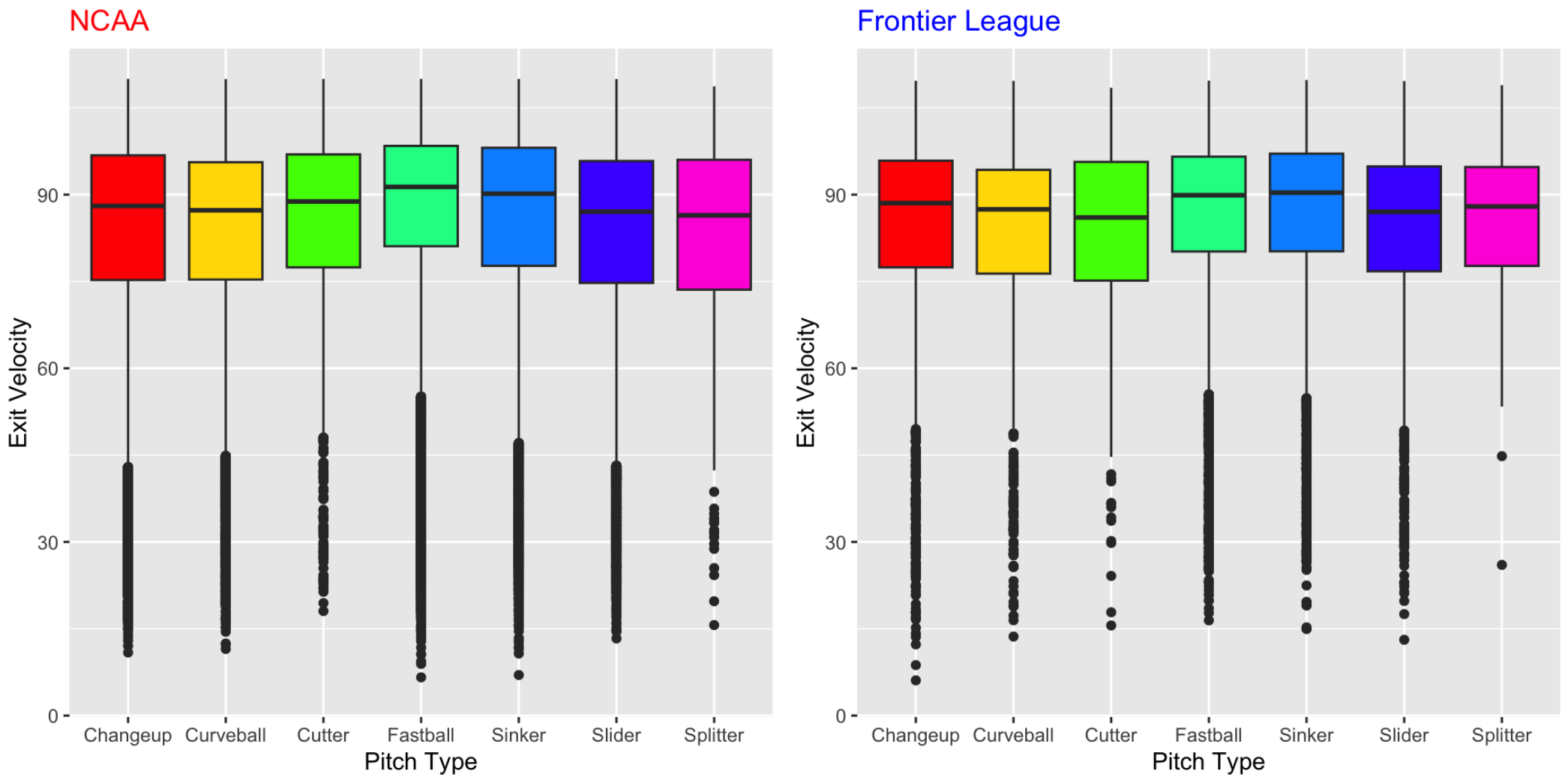
Average Pitch Velocity per inning by pitch type

Violin plot



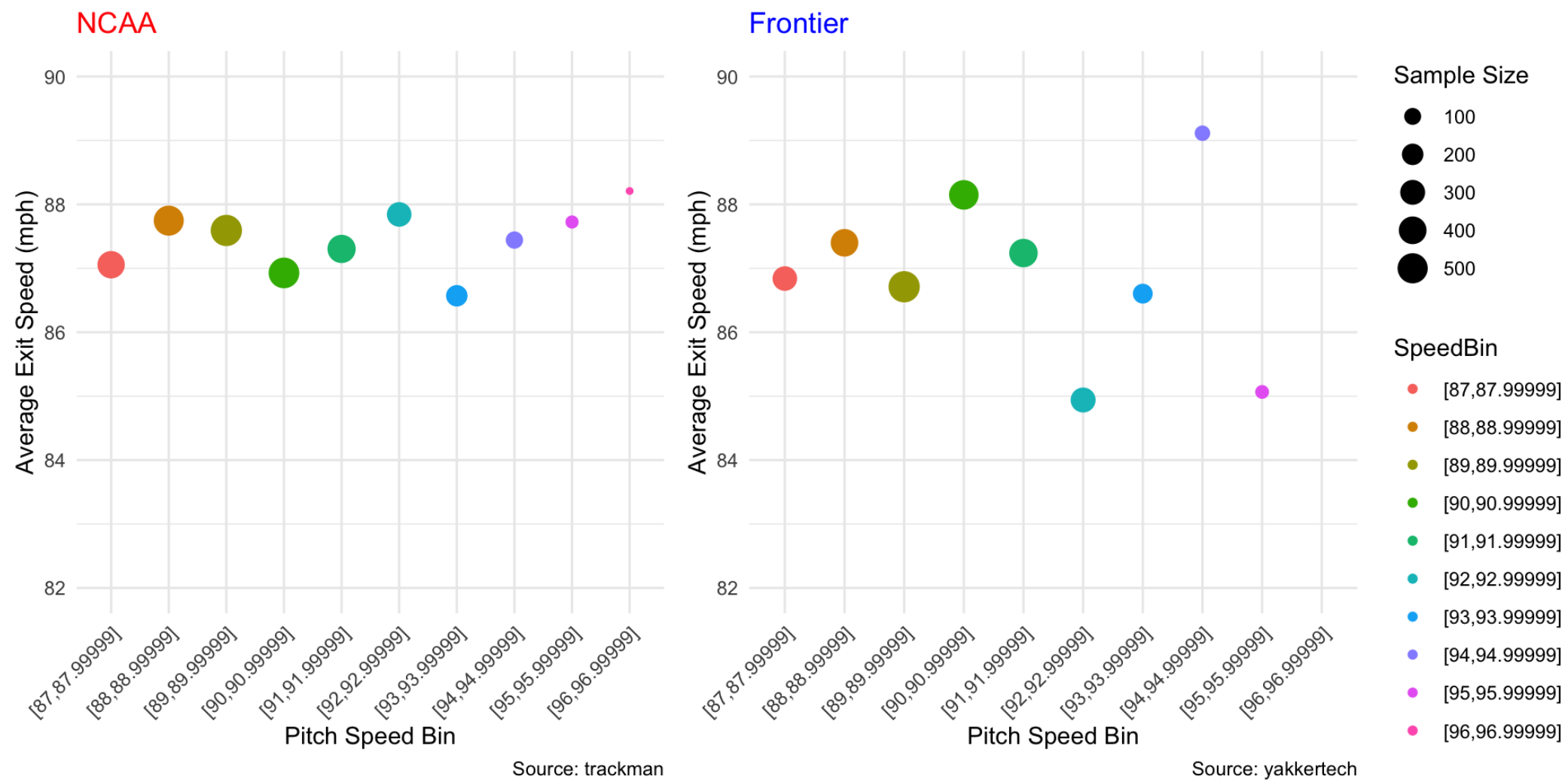
Violin plots of spin rates by pitch type

Exit speeds by pitch type



Box plots of exit speed by pitch type

Dead zones



Avg Exit Speed by Pitch Speed Bin of deadzone fastballs and sinkers

Killer plot: Situational Pitch Type Run Values and Exit Speeds

Choose Ball Count:

Any

Choose Strike Count:

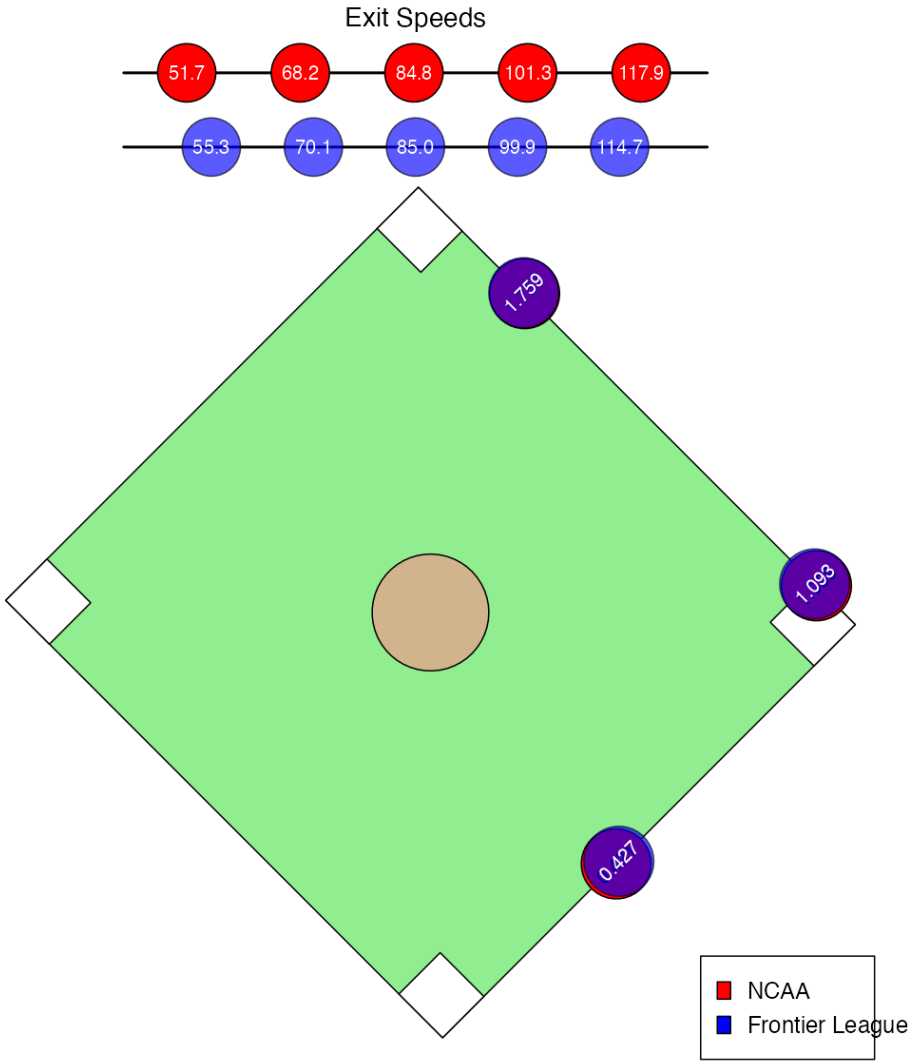
Any

Choose Pitcher Handedness:

Any

Choose Pitch Type:

Any



% of total balls in play (NCAA): 100%

% of total balls in play (Frontier League): 100%

NCAA impact of release speed on run values of hits

NCAA - Pitch Speed & Run value

	Dependent variable:				
	RunValue				
	Fastball	Curveball	Slider	Changeup	Sinker
	(1)	(2)	(3)	(4)	(5)
RelSpeed	0.005***	0.006***	0.005***	0.005***	0.005***
	(0.00003)	(0.0001)	(0.00005)	(0.00004)	(0.00004)
Observations	52,689	19,244	28,083	35,545	33,207
R2	0.279	0.289	0.274	0.284	0.297
Adjusted R2	0.279	0.289	0.274	0.284	0.297

Note: *p<0.1; **p<0.05; ***p<0.01

Frontier impact of release speed on run value of hits

Frontier - Pitch speed & Run value

	Dependent variable:				
	RunValue				
	Fastball	Curveball	Slider	Changeup	Sinker
	(1)	(2)	(3)	(4)	(5)
RelSpeed	0.005***	0.006***	0.005***	0.005***	0.005***
	(0.0001)	(0.0002)	(0.0001)	(0.0001)	(0.0001)
Observations	7,266	2,095	3,759	3,888	3,975
R2	0.285	0.309	0.277	0.302	0.298
Adjusted R2	0.285	0.309	0.276	0.302	0.297

Note: *p<0.1; **p<0.05; ***p<0.01

Conclusion

- The leagues are very similar in terms of pitch metrics and exit speeds of hits
- Trying to predict outcomes like exit velocities and run values of individual pitches is hard because the data is noisy
- Frontier League relievers/closers seem to throw more fastballs/sinkers and throw harder than Frontier League starters
- Trackman is a more accurate tracking system than Yakkertech
- NCAA hitters are better at hitting fastballs, cutters, and MLB-level deadzone pitches than Frontier League hitters
- In future analyses it would be useful to compare against top-tier leagues like MLB or Japan's NPB