Approach Angle Visuals Documentation

Reasoning

The purpose of the approach angle visual is to illustrate to a hitter what kind of approach angles they should expect to see from a pitcher's pitch arsenal, as well as how this pitcher compares to the rest of the league in each pitch type. This is run weekly for each affiliate's opponent and contains every rostered pitcher with at least 75 pitches thrown since the start of the previous season.

Building Models to Predict VAA/HAA

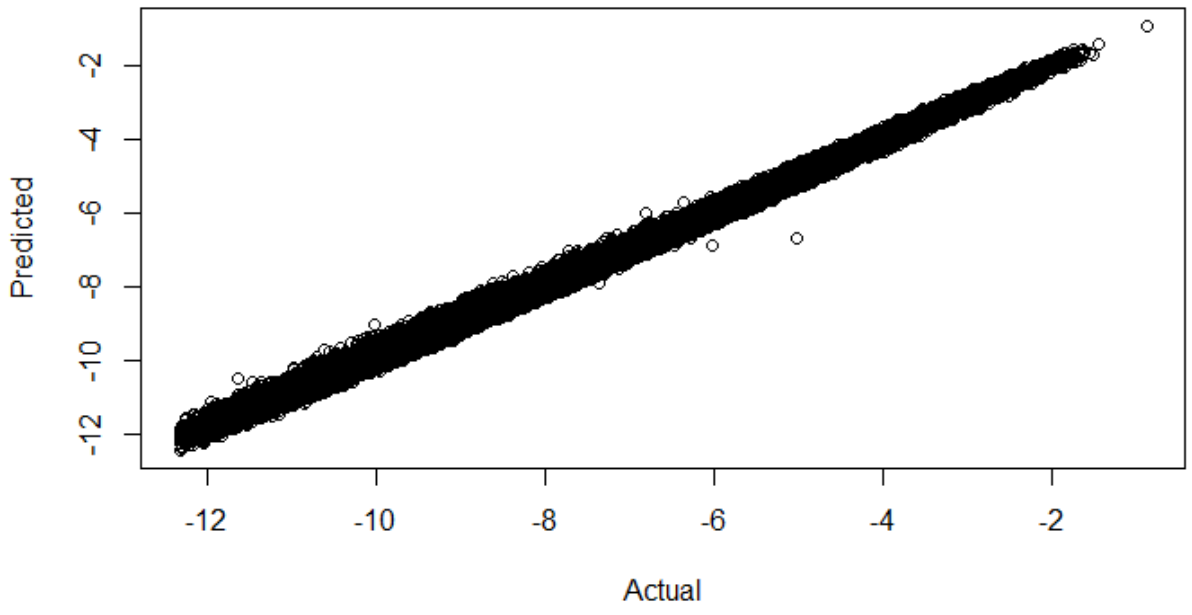
To start, pitch data was gathered from the 2023 MLB season which included pitch hand, pitch type, release speed, release coordinates, extension, pitch location, induced vertical (IVB) and horizontal (HB) break, and vertical approach (VAA) and horizontal approach (HAA) angle.

To build a model to predict VAA, the features selected were:

- Release Speed
- Z Release Coordinates
- Extension
- Z Pitch Location
- IVB

With VAA as the target

A multiple linear regression model was constructed and the results were excellent, returning an R^2 of 0.996 and an RMSE of 0.11. The predicted vs actual VAAs were plotted against one another in the graph below.



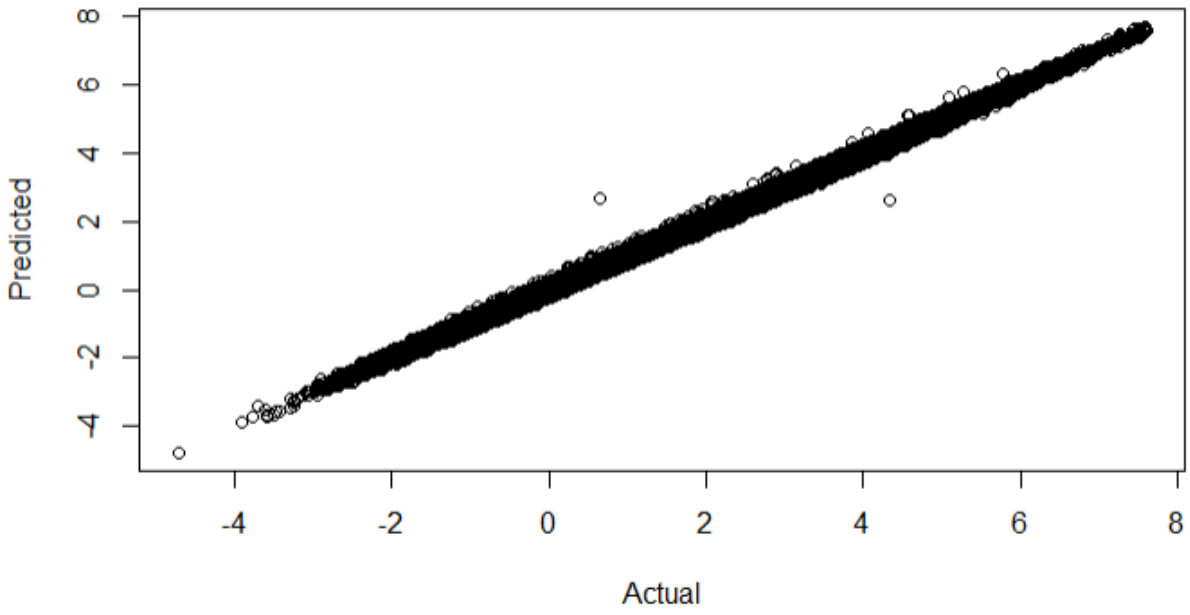
Approach Angle Visuals Documentation

Predicting HAA was done in a similar fashion, with the features selected being:

- Release Speed
- X Release Coordinates
- Extension
- X Pitch Location
- HB

With HAA as the target

The model for HAA also yielded excellent results, with an R^2 of 0.997 and and RMSE of 0.08. The predicted vs actual HAAs were plotted against each other in the graph below.



Predicting on Opposing Pitchers

Once the models were built, pitch data was acquired for the opponent's competition level from the start of the previous season to the current date. Because approach angles are affected by pitch location, all pitch X and Z location coordinates were adjusted to the middle of the strike zone to be location-neutral. From here, we predicted the new *adjusted VAA* and *adjusted HAA* using the linear models built on the '23 MLB data.

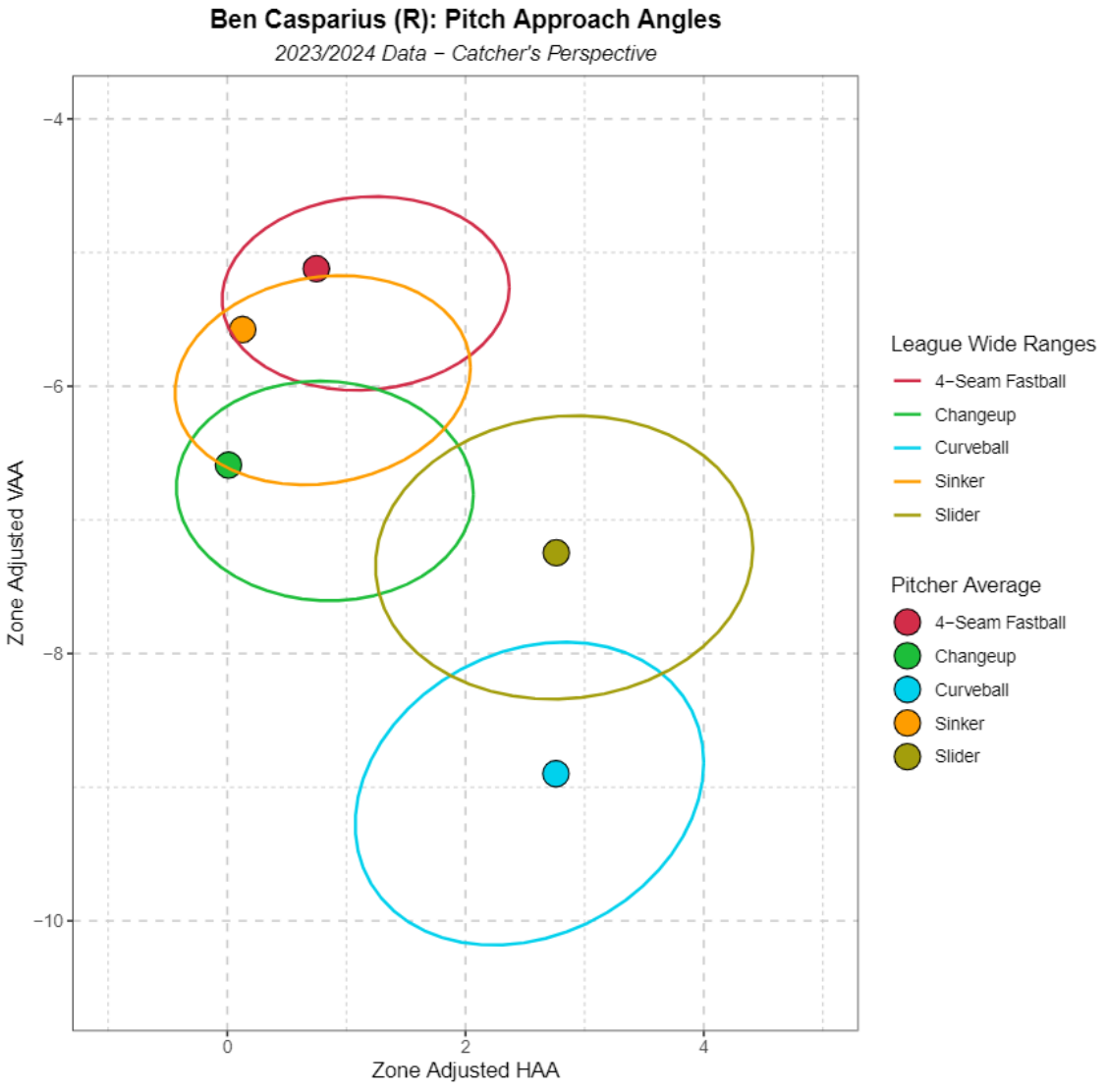
In a similar manner, pitch data for the currently rostered pitchers on the opposing team was gathered and predictions were made for adjusted VAA and HAA. The data was then grouped by pitcher and pitch type and mean VAAs and HAAs were found. For a pitch to be considered part of a pitcher's repertoire, he must have thrown it a minimum of 10 times.

Approach Angle Visuals Documentation

Making the Visual

Using ggplot, each pitcher's average adjusted VAA and HAA were plotted against one another, grouped by pitch type. For left-handed pitchers, the HAA was flipped across the y-axis for better interpretability. The overall competition level data was used to create 80% confidence ellipses which were plotted on the same visual. This gives the viewer an idea of how a pitcher's repertoire stacks up against other pitchers at the same level. For example, pitchers with an average 4-seam fastball VAA near the top or above the confidence ellipse likely have some of the most effective heaters because they get the "rising-effect" so many teams want and so many hitters hate.

Below is an example of how an approach angle visual might look for a pitcher at AAA.



Approach Angle Visuals Documentation

Message to Code Manager

The script for both the approach angle models and visuals can be found in Databricks in databricks_analytics/Approach_Angle_Visuals. You will need to set up a job to run each Sunday around 6 pm so each affiliate can receive the visuals before their designated off-day on Mondays.

Also, once the season gets going you can assess a point where there is enough data to remove the previous season's numbers and only use the current season's data. This will be a couple months into the season typically.