

## SJSU Baseball – First Report

**Goal 1: Find if there is a correlation between “HorzBreak and PFxx” and “VertBreak and PFxz”**

### Trackman Keyword Definitions:

**HorzBreak** – (Distance, inches) – between where the pitch actually crosses the front of home plate side-wise, and where it would have crossed home plate side-wise if had it traveled in a perfectly straight line from release.

Positive = the pitch was to the right from the pitcher’s perspective

Negative = the pitch was to the left from the pitcher’s perspective

**PFxx** – the horizontal (left-right) movement of the pitch during the last 40 feet before the front of home plate, as compared to a theoretical pitch thrown at the same speed with no spin-induced movement.

**VertBreak** – (Distance, inches) – between where the pitch actually crosses the front of home plate height-wise, and where it would have crossed home plate height-wise if had it traveled in a perfectly straight line from release, completely unaffected by gravity.

**PFxz** – the vertical (up-down) movement of the pitch during the last 40 feet before the front of home plate, as compared to a theoretical pitch thrown at the same speed with no spin-induced movement

\*\* It should be noted that when the word, “true” is referenced, it refers to the pitcher’s distinct contribution to the movement of the ball.

### Data information:

The data consist of four games (SJSU vs UOP 2.19.19, SJSU vs UCR 2.22.19, SJSU vs UCR 2.23.19, SJSU vs UCR 2.24.19), total 1303 observations. After removing the outliers and treating the missing values, the data has 1291 observations.

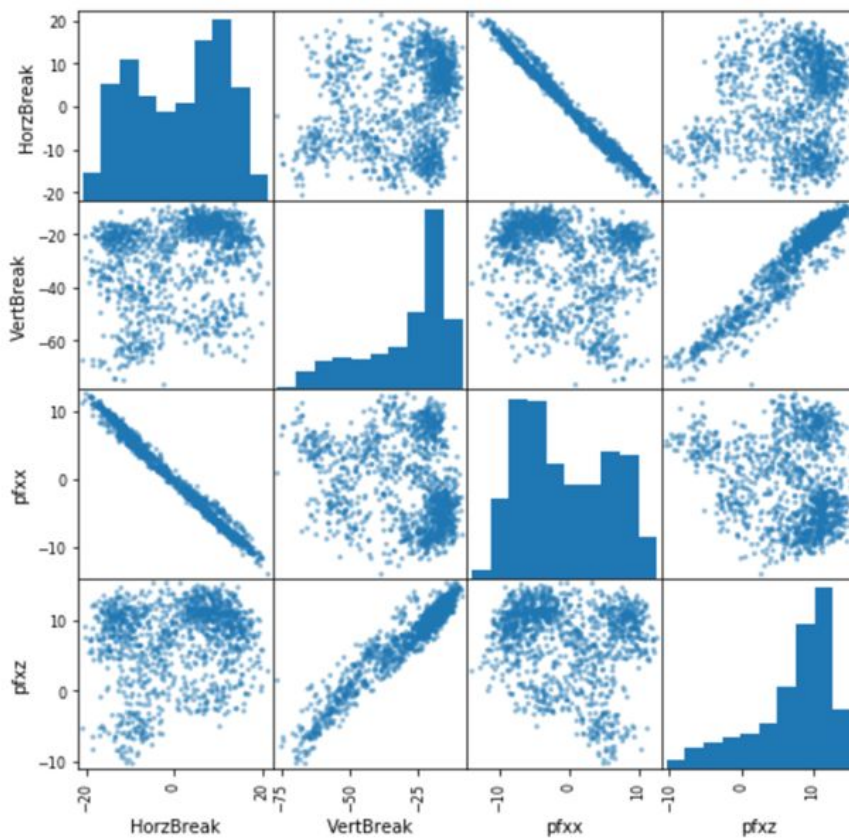
### Correlation matrix:

	<b>HorzBreak</b>	<b>VertBreak</b>	<b>pfx</b>	<b>pfxz</b>
<b>HorzBreak</b>	1.000000	0.239349	-0.993234	0.195656
<b>VertBreak</b>	0.239349	1.000000	-0.234776	0.957732
<b>pfx</b>	-0.993234	-0.234776	1.000000	-0.185633
<b>pfxz</b>	0.195656	0.957732	-0.185633	1.000000

We can see that the HorzBreak and pfx are highly correlated. VertBreak and pfxz are also highly correlated. HorzBreak and pfxz do not show much of correlation, neither do VertBreak and pfx.

Heat map:

A more distinct plot to show the correlations.

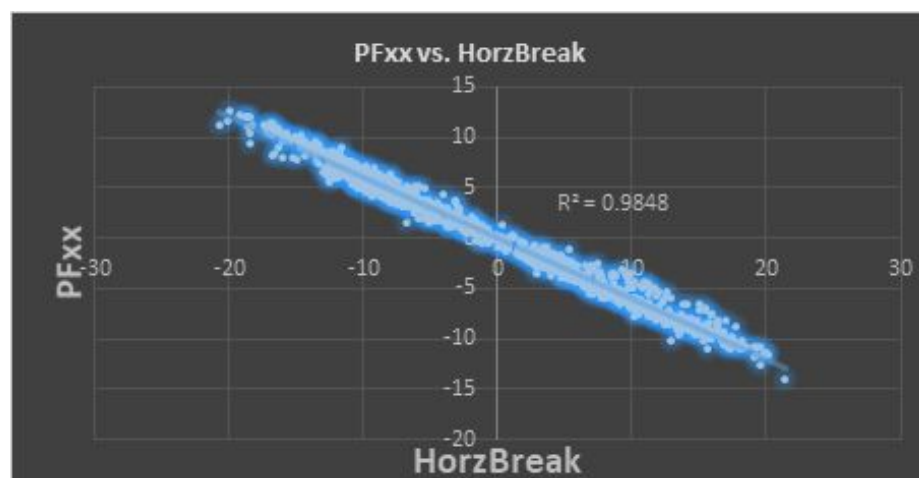
Scatter matrix:

The scatter matrix plot shows linear relationship between HorzBreak/pfx and VertBreak/pfxz. The well scattered plots and distribution shows that we have a sufficient sample size to perform data analysis.

$R^2$  table:

	HorzBreak	VertBreak	pfx	pfxz
HorzBreak	1.000000	0.057288	0.986513	0.038281
VertBreak	0.057288	1.000000	0.055120	0.917251
pfx	0.986513	0.055120	1.000000	0.034460
pfxz	0.038281	0.917251	0.034460	1.000000

### HorzBreak vs. PFxx

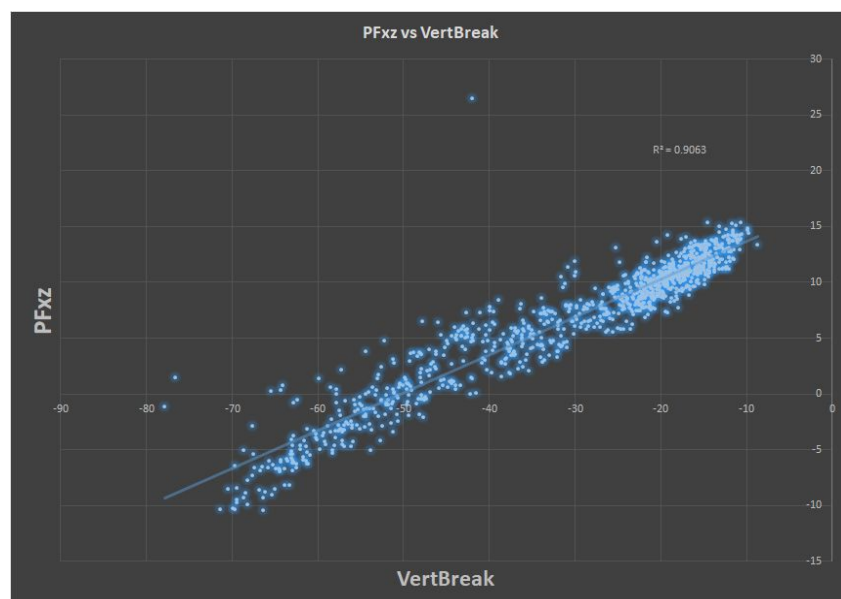


HorzBreak plotted against PFxx shows, “given the last 40 feet of ‘true’ movement (a same velocity thrown ball with no spin-‘induced’ might have some movement) a pitch had, how much movement did this ball have crossing the front plane of home plate compared to a similar thrown velocity ball if it had traveled in a straight line. We are of the assumption that PFxx is taken from the catcher’s perspective, meaning that a negative value means the pitch traveled to the left of the catcher and vice versa for a positive value.

Point of release should not be a factor in HorzBreak since the value is calculated by using the reference point of a similar straightly thrown ball from the same point of release.

With a  $R^2$  value of 0.9865 our analysis shows that there is a strong correlation between PFxx vs. HorzBreak. A possible interpretation of this result is that when a pitch thrown has an upper level of movement in the last 40 feet traveled (whether negative or positive) the ball will cross the front plane of home plate in that direction.

## VertBreak vs. PFxz



These two variables measure the vertical movement of a pitch. It shows: given the true vertical (up and down) movement of a pitch in the last 40 feet traveled, how far ‘down’ (compared to a ball on a straight line) did the ball cross the front plane of home plate compared to if had traveled in a perfectly straight line, unaffected by gravity from release.

Our analysis shows that there is a strong correlation between VertBreak and PFxz. The  $R^2$  value is 0.91725.

## **Goal 2: What are the averages & standard deviation for each pitch for all SJSU pitcher’s “metrics”?**

Please download the csv file:

<https://drive.google.com/open?id=1xMjqYFqDj7g37C2coAHjpIzBcwePBcxt>

The file contains the averages and standard deviation for each pitcher with each pitch type.

## **Goal 3: Do any of the metrics change when changing between the “Windup” and the “Stretch”? If so by how much?**

Please download the csv file:

<https://drive.google.com/open?id=1JEmrBwDjDB580gnt7bakaFCRJhv8gcfl>

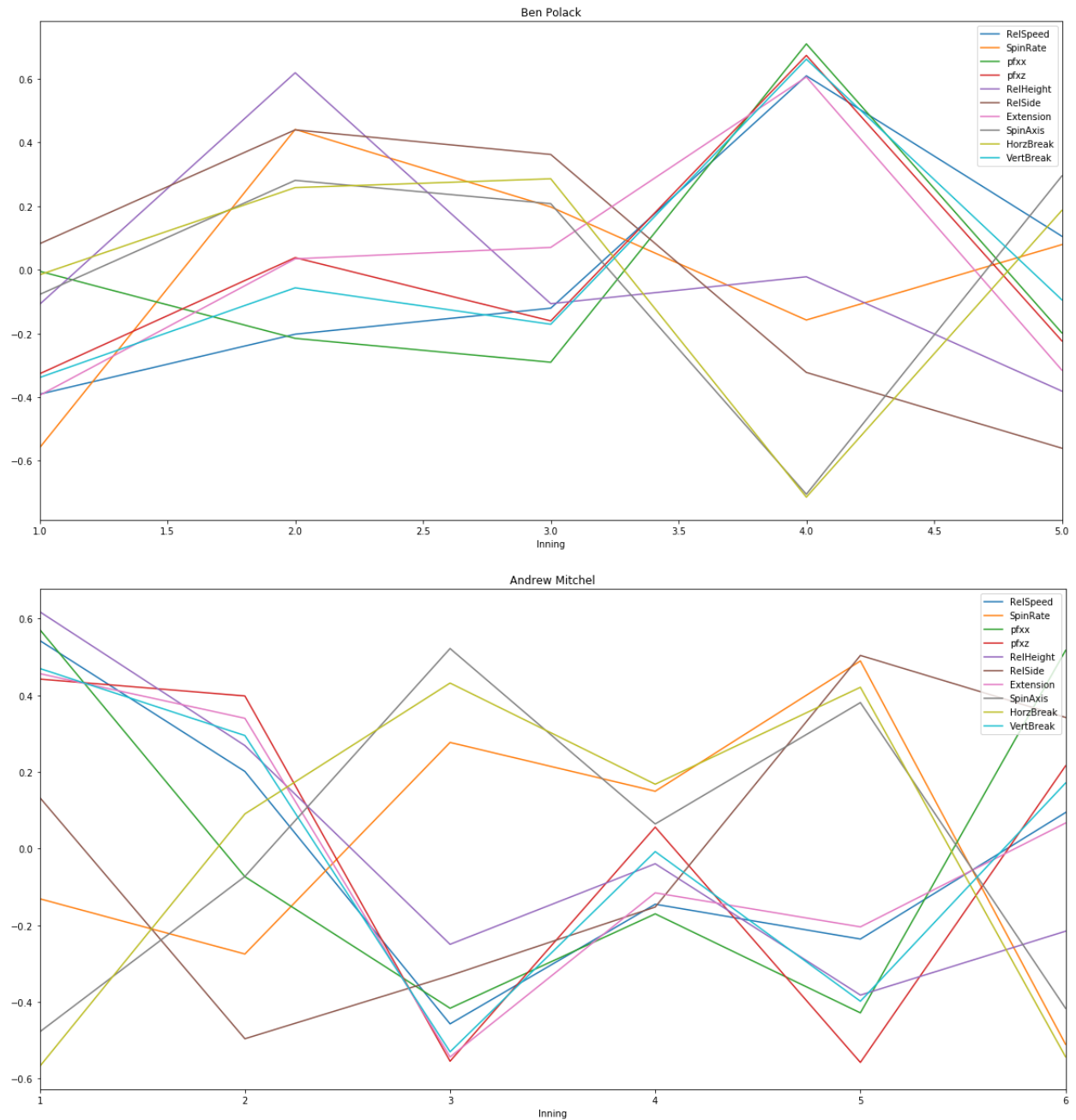
The file contains the comparison of all the metrics between the pitcher set “Windup” and “Stretch”. The values represent mean(“metric in windup”) - mean(“metric in stretch”). For example:

The first row of the file is:

	RelSpeed	SpinRate	pfx	pfxz	RelHeight	RelSide	Extension	SpinAxis	HorzBreak	VertBreak
Bateson, Tre	4.58632636	-105.29177	-4.5743761	4.04239329	0.06095524	-0.0220866	-0.0260343	35.5361759	8.02533653	12.7310072

The number in column “RelSpeed” means the mean of the RelSpeed in Windup minus the mean of the RelSpeed in Stretch for player Trevor Bateson. The same mechanism applies to all metrics.

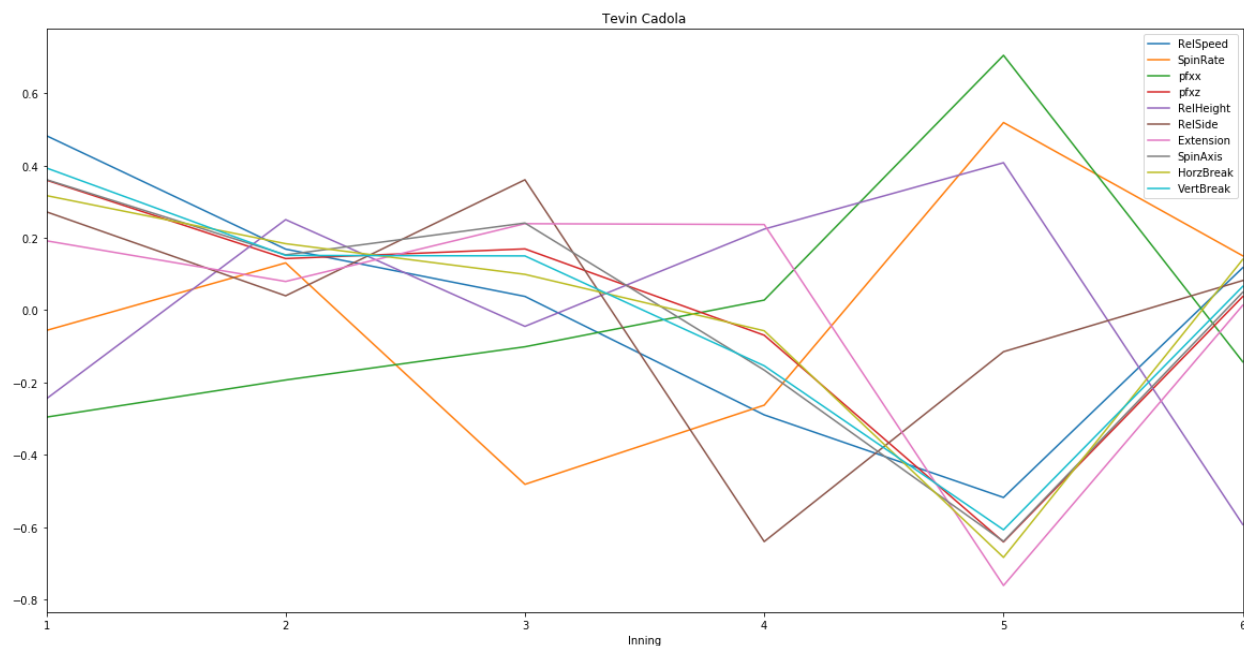
**Goal 4: For Polack, Mitchel & Cadola (last names) how did the “metrics” fluctuate as they pitched more innings?**



Devin Humm

George Qiao

2/28/2019



**Goal 4: For each SJSU pitcher, how often was each pitch thrown per inning?**

Please download the file:

<https://drive.google.com/open?id=1IVJsNUIPa3ywFjXjeayeBV61h9nu86>