EXPLORING PITCH DATA IN R

Batted ball outcomes - contact rate

Successful pitching characteristics

- How often do batters make contact on various pitches?
- How does location impact contact rate?
- Most effective two-strike pitches?
- How can we visualize how batters swing?

Pitch outcomes: Contact rate

- **Contact rate:** percentage of swings by batter that do not miss
- **Whiff rate:** percentage of swings by batter that miss

The `pitch_result` variable

```
> levels(greinke_sub$pitch_result)
 [1] "Ball"                      "Ball In Dirt"
 [3] "Called Strike"             "Foul"
 [5] "Foul (Runner Going)"       "Foul Bunt"
 [7] "Foul Tip"                  "Hit By Pitch"
 [9] "In play, no out"           "In play, out(s)"
[11] "In play, run(s)"           "Missed Bunt"
[13] "Swinging Strike"           "Swinging Strike (Blocked)”

> head(greinke_sub$pitch_result)
[1] Ball               Swinging Strike   Called Strike
[4] Swinging Strike    Swinging Strike   Swinging Strike
```

Batted balls

```
> head(play_out[, 5:9])
      pitch_result atbat_result start_speed    z0     x0
13 In play, out(s)      Pop Out         87.1 6.224 -0.904
14 In play, out(s)     Sac Bunt         90.0 5.987 -1.185
42 In play, out(s)    Groundout         85.5 6.132 -1.013
45 In play, out(s)    Groundout         91.7 6.053 -1.112
52 In play, out(s)     Forceout         86.4 6.463 -0.613
64 In play, out(s)    Groundout         94.2 6.124 -0.736

> head(play_no_out[, 5:9])
       pitch_result atbat_result start_speed    z0     x0
43  In play, no out       Single         87.4 6.213 -1.060
65  In play, no out       Single         91.7 6.387 -0.838
155 In play, no out       Double         77.4 6.519 -0.924
156 In play, no out       Single         90.9 5.949 -0.989
157 In play, no out       Single         93.0 6.178 -0.709
199 In play, no out       Single         85.4 6.435 -1.061
```

EXPLORING PITCH DATA IN R

Let's practice!

EXPLORING PITCH DATA IN R

Using `ggplot2`

Wide and long format

```
> head(wide) Multiple measurements per row
   zone   px  pz exit_speed_rhb exit_speed_lhb
1     1 -1.5 4.5            NA            NA
2     2 -0.5 4.5            NA          72.0
3     3  0.5 4.5          92.0            NA
4     4  1.5 4.5            NA            NA
5     5 -1.5 3.5          60.5          91.0
6     6 -0.5 3.5          88.0          91.0

> long[15:20, ] One measurement per row
   zone   px  pz batter_stand exit_speed
15   15  0.5 1.5          RHB       89.8
16   16  1.5 1.5          RHB       84.2
17   17 -1.5 0.5          RHB         NA
18   18 -0.5 0.5          RHB       89.0
19   19  0.5 0.5          RHB       84.6
20   20  1.5 0.5          RHB       91.1
```

Locational coordinates: `locgrid`

```
> head(locgrid)
  zone   px  pz
1    1 -1.5 4.5
2    2 -0.5 4.5
3    3  0.5 4.5
4    4  1.5 4.5
5    5 -1.5 3.5
6    6 -0.5 3.5

> tail(locgrid)
   zone   px  pz
15   15  0.5 1.5
16   16  1.5 1.5
17   17 -1.5 0.5
18   18 -0.5 0.5
19   19  0.5 0.5
20   20  1.5 0.5
```

`ggplot2`: Layers and plotting

- **`ggplot()`**: plotting
- **`aes()`**: aesthetics
- **`ggtitle()`**: figure titles
- **`labels()`**: axis labels
- **`theme()`**: adjust specifics for axes, etc.
- **`geom_tile()`**: for making a grid
- **`scale_fill_gradientn()`**: filling grids with color
- **`facet_grid()`**: side by side plotting windows
- **`annotate()`**: writing text

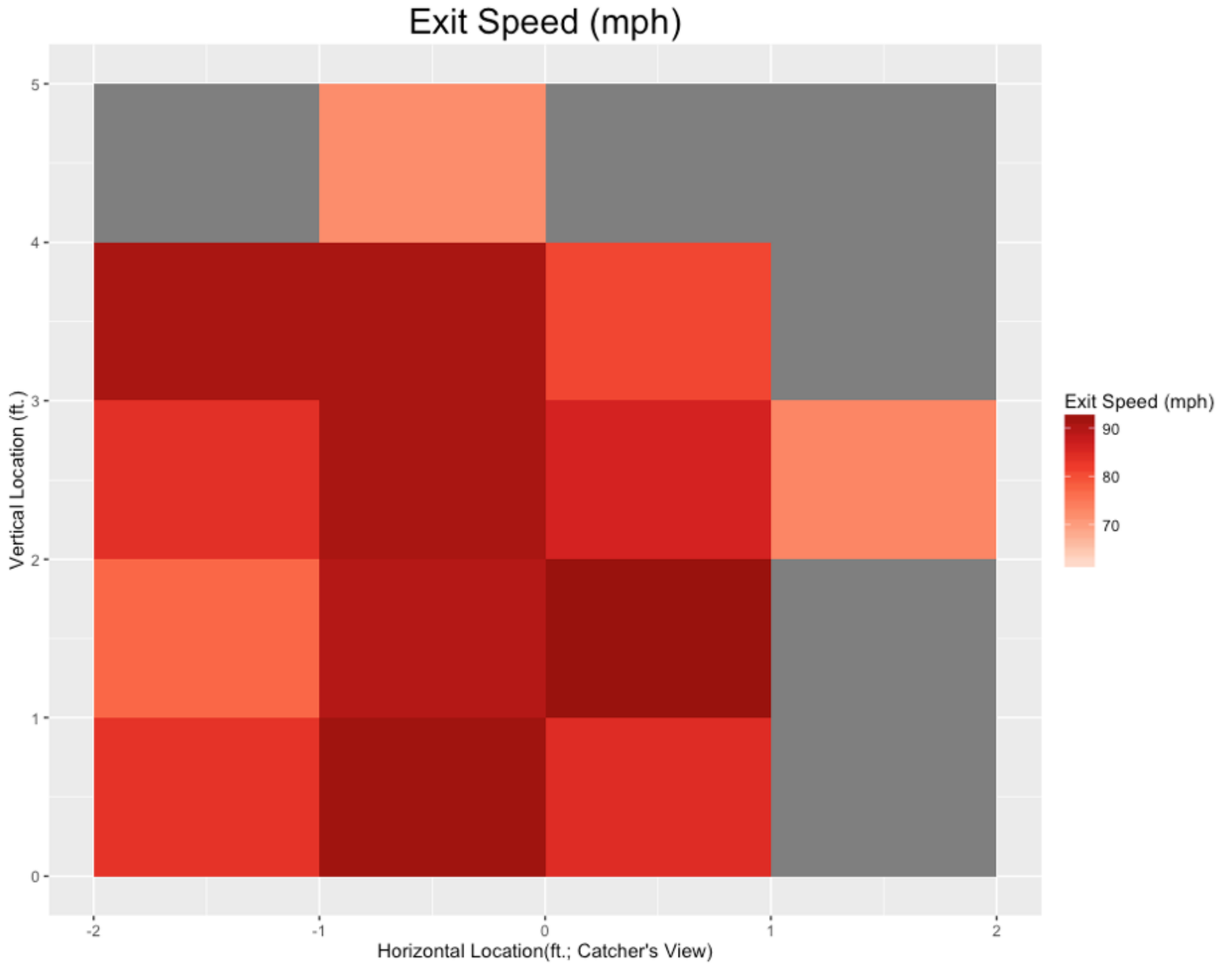
Beautiful plots

```
> library(ggplot2)
> library(RColorBrewer)

> plot_exit <- ggplot(data = exit_tidy, aes(x = px, y = pz)) +
    geom_tile(data = exit_tidy, aes(fill = exit_speed)) +
    scale_fill_gradientn(name = "Exit Speed (mph)",
                         colours = c(brewer.pal(n = 7, name = "Reds"))) +
    ggtitle("Exit Speed (mph)") +
    labs(x = "Horizontal Location(ft.; Catcher's View)",
         y = "Vertical Location (ft.)") +
    theme(plot.title = element_text(size = 20))

> plot_exit
```

Beautiful plots



EXPLORING PITCH DATA IN R

Let's practice!

EXPLORING PITCH DATA IN R

Batted ball outcomes - exit velocity

Pitch outcomes: Exit velocity

```
> head(swings$batted_ball_velocity)
[1] 104  NA  NA  NA  NA  NA

> sort(tapply(swings$batted_ball_velocity, swings$atbat_result, mean))
Fielders Choice Out    Fan interference     Fielders Choice
           61.00000             65.00000             65.00000
        Double Play              Sac Fly               Double
           81.00000             93.00000             96.53488
             Triple             Home Run
          100.50000            104.07143
```

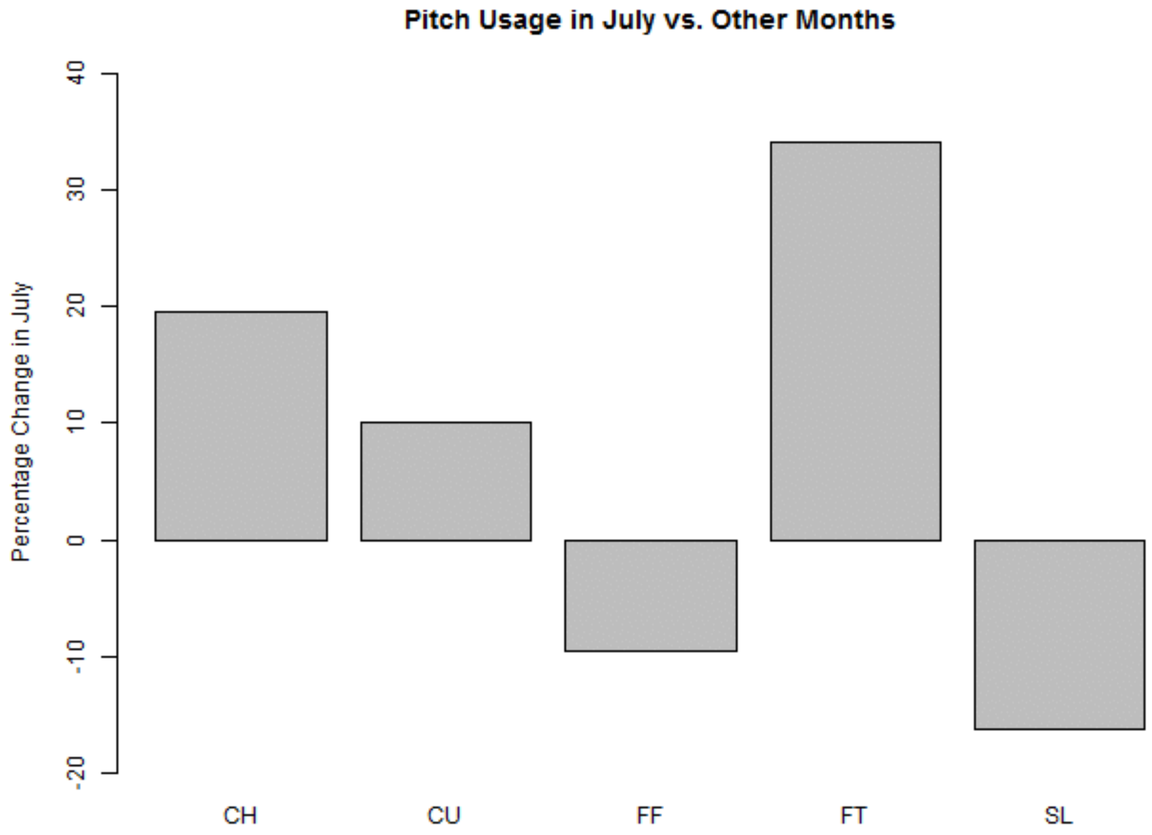
EXPLORING PITCH DATA IN R

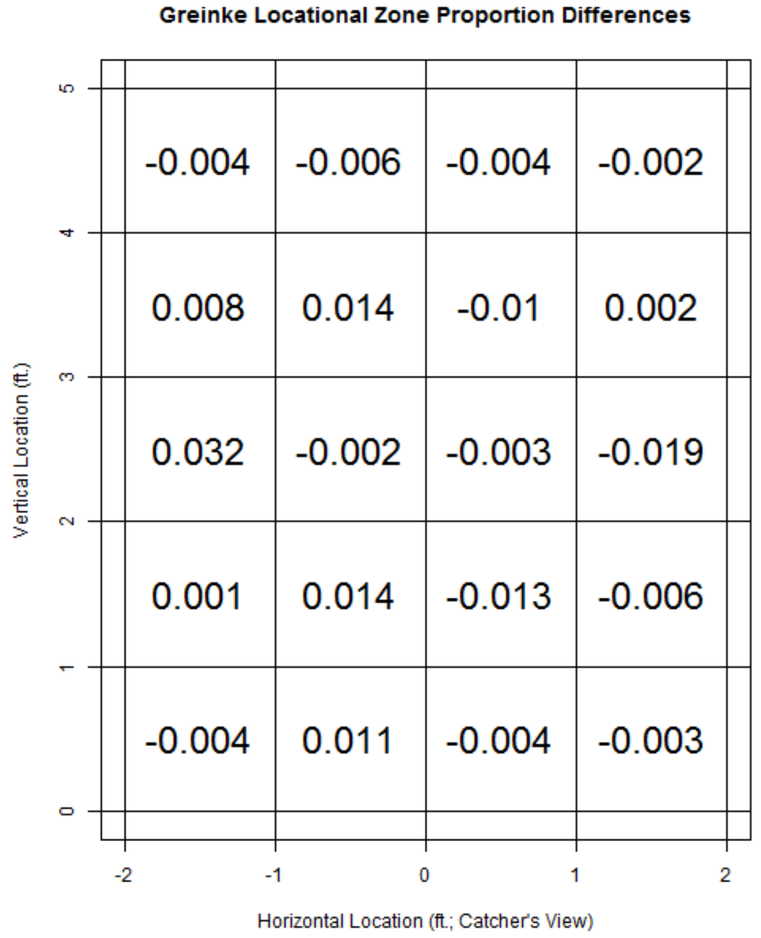
Let's practice!

EXPLORING PITCH DATA IN R

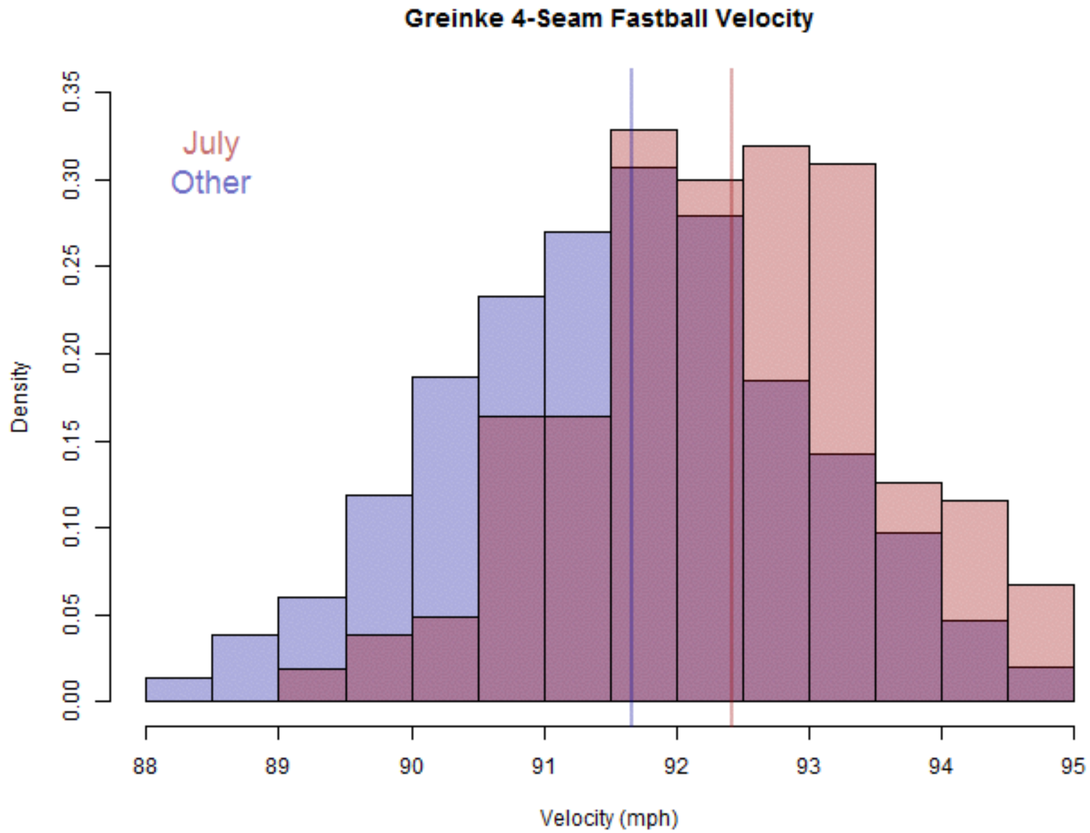
Wrap-up

Summary of findings





Summary of findings

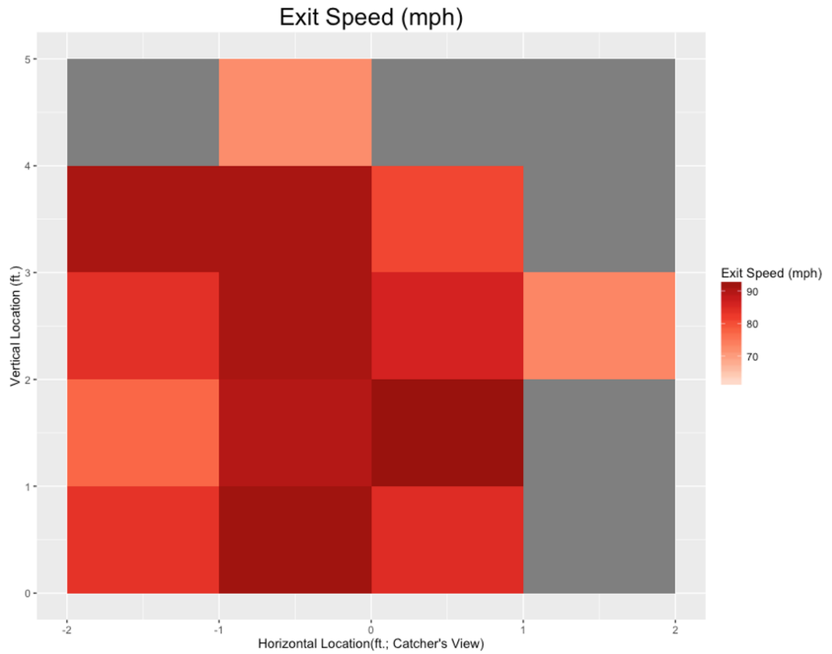


Fangraphs 2015 pitcher leaderboards

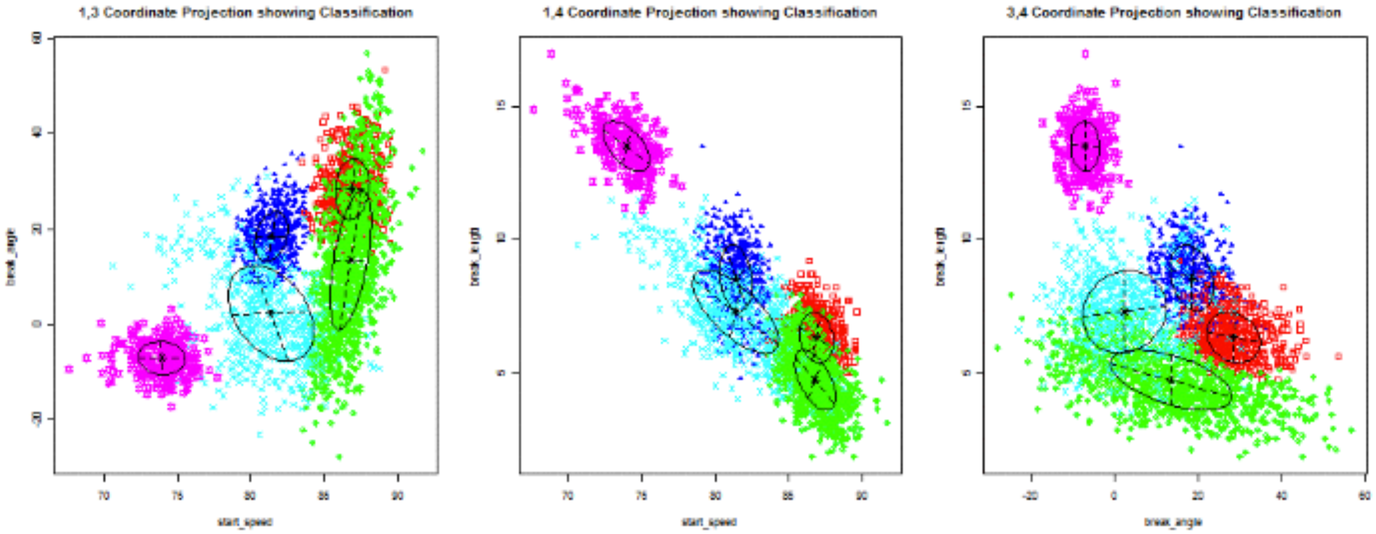
#	Name	Team	W	L	SV	G	GS	IP	K/9	BB/9	HR/9	BABIP	LOB%	GB%	HR/FB	ERA	FIP	xFIP	WAR
1	Clayton Kershaw	Dodgers	16	7	0	33	33	232.2	11.64	1.62	0.58	.281	78.3 %	50.0 %	10.1 %	2.13	1.99	2.09	8.6
2	Jake Arrieta	Cubs	22	6	0	33	33	229.0	9.28	1.89	0.39	.246	80.0 %	56.2 %	7.8 %	1.77	2.35	2.61	7.3
3	David Price	- - -	18	5	0	32	32	220.1	9.19	1.92	0.69	.290	78.6 %	40.4 %	7.8 %	2.45	2.78	3.24	6.4
4	Max Scherzer	Nationals	14	12	0	33	33	228.2	10.86	1.34	1.06	.268	79.6 %	35.8 %	10.5 %	2.79	2.77	2.88	6.4
5	Chris Sale	White Sox	13	11	0	31	31	208.2	11.82	1.81	0.99	.323	73.2 %	42.6 %	12.5 %	3.41	2.73	2.60	6.2
6	Dallas Keuchel	Astros	20	8	0	33	33	232.0	8.38	1.98	0.66	.269	79.4 %	61.7 %	13.6 %	2.48	2.91	2.75	6.1
7	Zack Greinke	Dodgers	19	3	0	32	32	222.2	8.08	1.62	0.57	.229	86.5 %	48.0 %	7.3 %	1.66	2.76	3.22	5.9
8	Corey Kluber	Indians	9	16	0	32	32	222.0	9.93	1.82	0.89	.297	71.4 %	42.4 %	10.7 %	3.49	2.97	3.05	5.5
9	Gerrit Cole	Pirates	19	8	0	32	32	208.0	8.74	1.90	0.48	.304	74.9 %	48.0 %	6.5 %	2.60	2.66	3.16	5.4
10	Chris Archer	Rays	12	13	0	34	34	212.0	10.70	2.80	0.81	.295	73.1 %	46.1 %	10.4 %	3.23	2.90	3.01	5.3

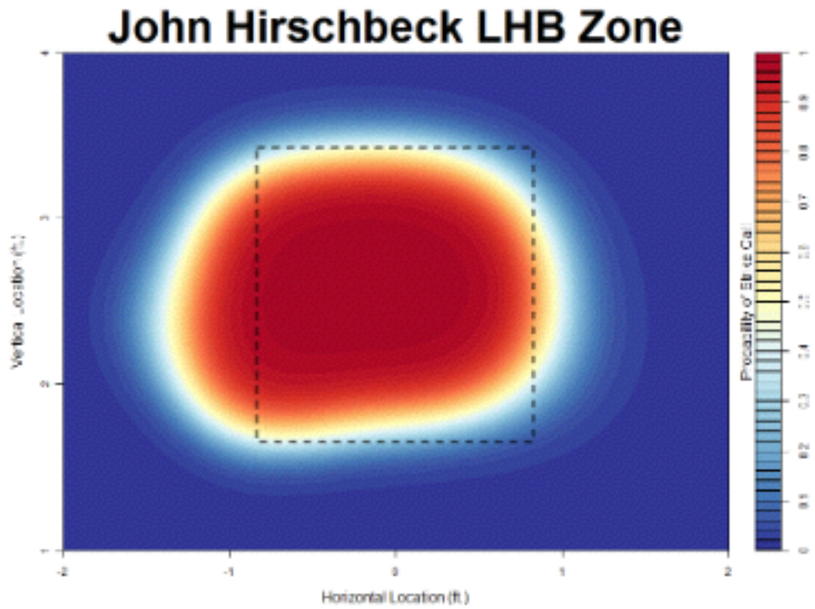
Course progression

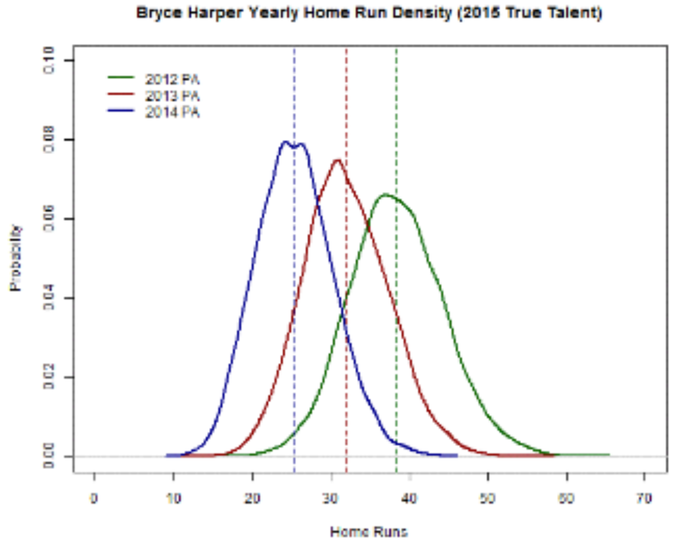
	p_name	pitcher_id	batter_stand	pitch_type	pitch_result	atbat_result	start_speed	z0	x0	pfx_x	pfx_z	px	pz	break_angle	break_length
1	Zack Greinke	425844	R	FF	Ball	Walk	94.2	5.997	-0.675	-4.457	9.760	1.714	1.925	24.8	3.5
2	Zack Greinke	425844	R	FF	Swinging Strike	Single	92.4	6.281	-0.760	-1.590	11.400	0.589	3.271	10.1	2.7
3	Zack Greinke	425844	R	FF	Called Strike	Home Run	92.7	6.168	-0.958	-1.884	9.245	0.399	2.918	9.2	3.5
4	Zack Greinke	425844	R	SL	Swinging Strike	Strikeout	86.9	6.077	-0.939	3.594	0.762	0.764	1.306	-11.4	8.0
5	Zack Greinke	425844	R	FF	Swinging Strike	Strikeout	92.8	6.107	-0.524	-0.558	11.134	1.517	2.193	-0.4	2.8
6	Zack Greinke	425844	R	SL	Swinging Strike	Strikeout	87.8	6.321	-0.948	4.313	0.132	0.695	3.431	-13.6	7.8
7	Zack Greinke	425844	R	CH	Swinging Strike (Blocked)	Strikeout	90.3	5.865	-1.252	-7.525	1.745	-0.131	0.922	22.5	7.4
8	Zack Greinke	425844	R	FF	Called Strike	Strikeout	92.7	6.103	-0.752	-4.795	10.161	0.994	2.239	25.1	3.8
9	Zack Greinke	425844	R	SL	Foul	Groundout	85.5	6.221	-0.983	2.285	2.247	0.873	2.344	-8.4	7.5
10	Zack Greinke	425844	R	SL	Swinging Strike	Walk	87.3	6.043	-0.951	3.215	1.771	1.125	1.292	-11.3	7.4
11	Zack Greinke	425844	R	SL	Swinging Strike	Strikeout	87.4	6.312	-0.975	1.309	3.610	1.181	2.358	-6.5	6.7
12	Zack Greinke	425844	R	FF	In play, run(s)	Home Run	92.3	6.180	-1.206	-2.761	9.134	-0.165	3.503	15.8	3.5
13	Zack Greinke	425844	R	SL	In play, out(s)	Pop Out	87.1	6.224	-0.904	1.390	0.847	0.075	1.689	-4.8	7.8
14	Zack Greinke	425844	R	FF	In play, out(s)	Sac Bunt	90.0	5.987	-1.185	-3.643	10.472	-0.453	2.730	21.5	3.6
15	Zack Greinke	425844	R	FF	Swinging Strike	Strikeout	93.0	5.972	-0.858	-2.308	10.355	0.587	2.567	13.5	3.1
16	Zack Greinke	425844	R	FF	Ball	Single	91.8	6.121	-0.661	-1.507	9.496	1.584	2.695	5.8	3.5
17	Zack Greinke	425844	R	CH	Ball	Strikeout	89.9	5.919	-1.068	-9.532	1.380	0.671	0.536	26.1	8.0
18	Zack Greinke	425844	R	FT	Ball	Strikeout	90.0	5.939	-1.046	-8.678	2.809	0.370	1.460	26.7	7.2
19	Zack Greinke	425844	R	FF	Ball	Strikeout	93.8	6.219	-0.711	-2.753	10.886	1.098	3.352	17.7	2.8
20	Zack Greinke	425844	R	FF	Called Strike	Groundout	91.8	6.146	-0.736	-1.660	8.780	1.269	2.459	5.6	4.0
21	Zack Greinke	425844	R	CH	Called Strike	Pop Out	89.5	6.024	-1.096	-8.325	1.235	0.705	1.561	23.0	7.8
22	Zack Greinke	425844	R	FF	Foul	Pop Out	92.4	6.273	-0.789	-1.816	11.093	-0.285	3.189	12.8	2.9



Expand and explore







EXPLORING PITCH DATA IN R

Thanks!