

イチローと松井、どっちが凄い？

Rでセイバーメトリクス

@gg_hatano

発表の目的

1. セイバーメトリクスに親しみたい(僕が)
2. Rで野球データを触れることを伝えたい
3. 作ったパッケージの宣伝をしたい

目次

- ・ Rと野球データ
- ・ 打撃能力の比較1: 得点創出能力XR
- ・ 打撃能力の比較2: 野球は1人でもできる
- ・ まとめ

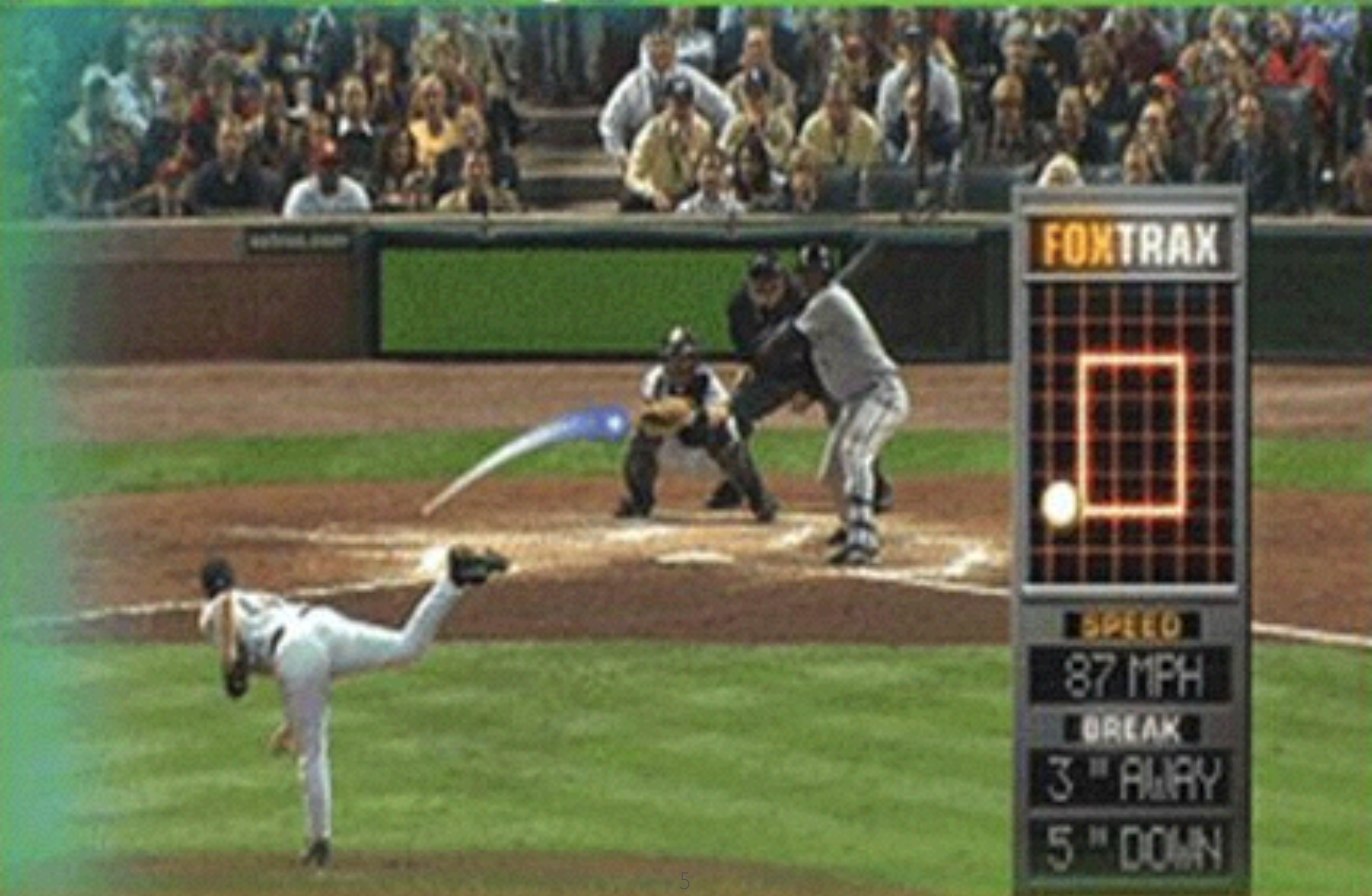
目次

Rと野球データ

- ・ 打撃能力の比較1: 得点創出能力XR
- ・ 打撃能力の比較2: 野球は1人でもできる
- ・ まとめ

PITCHf/x

FOX



Pitch f/xとは

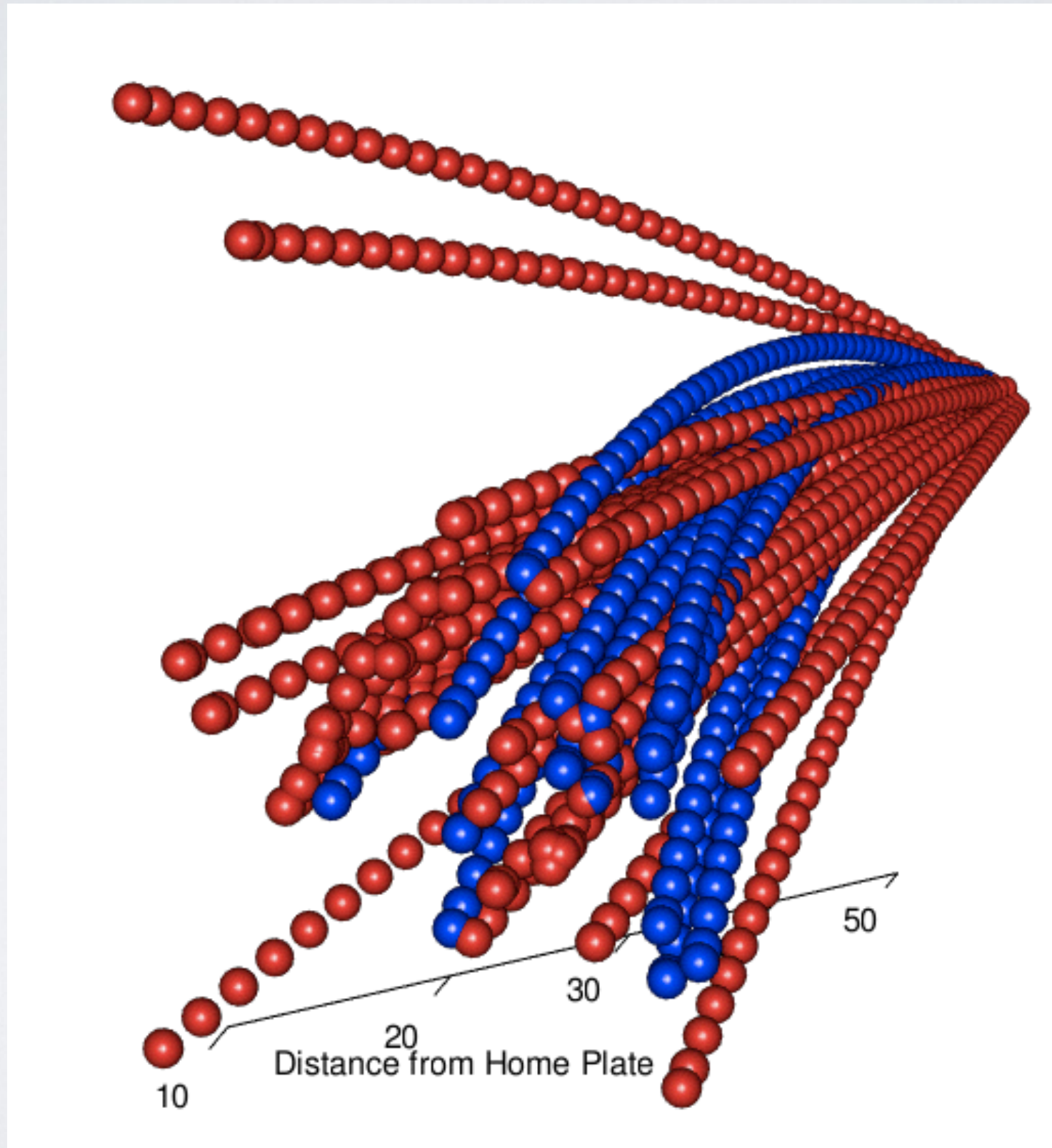
- ・ メジャーリーグ 全投球の軌道追跡システム
- ・ 初期位置, 初速, 回転, etc... (凄い)
- ・ すぐにデータが欲しい → `library(pitchRx)`

pitchRxで野球データ解析

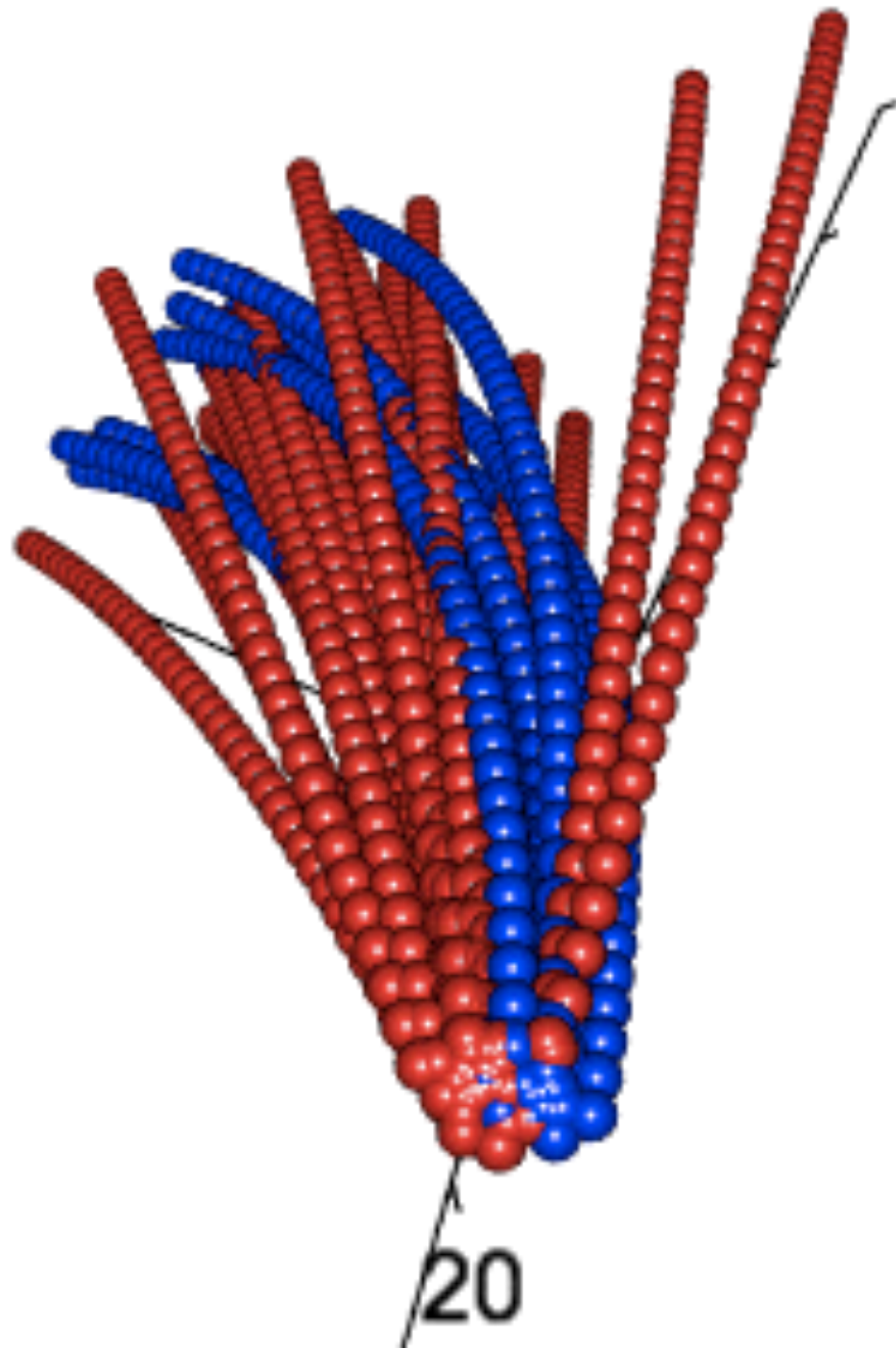
- ・ ダルビッシュの投球軌道 3D描画

<http://www7275uo.sakura.ne.jp/webGL/>

pitchRxで野球データ解析



pitchRxで野球データ解析



野球データ 分析の例

- ・ 初球の重要性を調べた

<http://gg-hogehoge.hatenablog.com/entry/2014/01/27/224946>

- ・ 投球テンポと援護率の関係を調べた

<http://gg-hogehoge.hatenablog.com/entry/2014/01/17/230927>

- ・ スピンの効いた球は打たれない？

<http://gg-hogehoge.hatenablog.com/entry/2014/03/07/035245>

Rと野球のデータ

- ・ 投球、打席、シーズン成績 全部揃っている

	投球	打席	シーズン
データ	pitch f/x	retrosheet	Lahman
Rのパッケージ	pitchRx	retrosheet	Lahman

Rと野球のデータ

- ・ 投球、打席、シーズン成績 全部揃っている

	投球	打席	シーズン
データ	pitch f/x	retrosheet	Lahman
Rのパッケージ	pitchRx	retrosheet	Lahman



作りました

Retrosheetとは

- ・ メジャーリーグ 全試合の全イベント記録
- ・ 約80年分のデータが無償公開
- ・ すぐにデータが欲しい → `library(retrosheet)`

Retrosheetとは

- ・ メジャーリーグ 全試合の全イベント記録
- ・ 約80年分のデータが無償公開
- ・ すぐにデータが欲しい → `library(retrosheet)`



作りました!!

Retrosheetでデータ分析

実況 「三者凡退でリズムを作りました!」

解説 「これは援護点が期待できますね」

Retrosheetでデータ分析

実況 「三者凡退でリズムを作りました!」

解説 「これは援護点が期待できますね」

Retrosheetでデータ分析

実況 「三者凡退でリズムを作りました!」

解説 「これは援護点が期待できますね」

Retrosheetでデータ分析

実況 「三者凡退でリズムを作りました!」

解説 「これは援護点が期待できますね」

=> 期待できるの??

Retrosheetでデータ分析

実況 「三者凡退でリズムを作りました!」

解説 「これは援護点が期待できますね」

=> 約2400試合のデータを使って、調べました

https://rpubs.com/gg_hatano/19465

三者凡退と援護点

- 。 分析の手順

1. 2013年の全試合データ

2. 投手が“三者凡退”で抑えたか判定

3. 次の攻撃で、得点が入ったかどうか

三者凡退と援護点

得点出来たかどうかの確率も見てみましょう。

```
```{r}
dat_sanbon_score %>%
 group_by(after_sanbon) %>%
 summarise("得点確率" = round(mean(score>0), 3)) %>%
 as.data.frame %>%
 mutate("三凡後フラグ" = after_sanbon) %>%
 select(3,2)
```
```

https://rpubs.com/gg_hatano/19465

三者凡退と援護点 得点確率

- 分析の結果：

| 三凡後フラグ 得点確率 | | |
|-------------|-------|-------|
| 1 | FALSE | 0.258 |
| 2 | TRUE | 0.250 |

三者凡退と援護点 得点確率

- 。 非三者凡退 => 次攻撃の得点確率: 25.8%

三凡後フラグ 得点確率

| | | |
|---|-------|-------|
| 1 | FALSE | 0.258 |
|---|-------|-------|

| | | |
|---|------|-------|
| 2 | TRUE | 0.250 |
|---|------|-------|

三者凡退と援護点 得点確率

- 三者凡退 => 次攻撃の得点確率: 25.0%

三凡後フラグ 得点確率

| | | |
|---|-------|-------|
| 1 | FALSE | 0.258 |
|---|-------|-------|

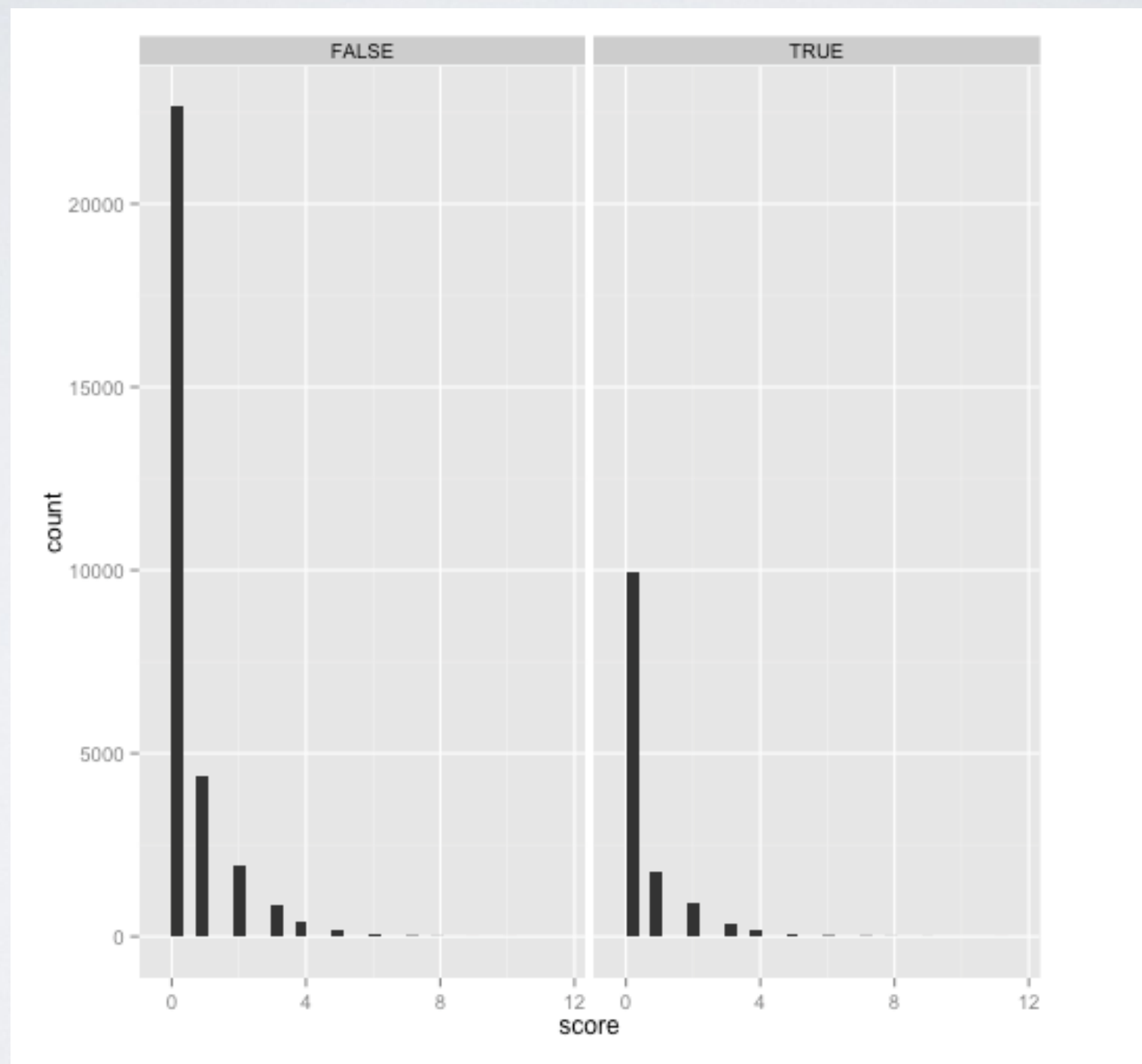
| | | |
|---|------|-------|
| 2 | TRUE | 0.250 |
|---|------|-------|

三者凡退と援護点 得点確率

- 。 分析の結果：得点確率は変わらない

| 三凡後フラグ 得点確率 | | |
|-------------|-------|-------|
| 1 | FALSE | 0.258 |
| 2 | TRUE | 0.250 |

三者凡退と援護点 得点分布



https://rpubs.com/gg_hatano/19465

Retrosheetでデータ分析

実況 「三者凡退でリズムを作りました!」

解説 「これは援護点が期待できますね」

Retrosheetでデータ分析

実況 「三者凡退でリズムを作りました!」

解説 「これは援護点が期待できますね」

=> そうでもないと思う

https://rpubs.com/gg_hatano/19465

野球データ 分析の例

- ・ 巨人の勝率が8割になる方法

<http://gg-hogehoge.hatenablog.com/entry/2014/08/20/203431>

- ・ 「やられたら、やりかえす」

<http://gg-hogehoge.hatenablog.com/entry/2014/06/12/193142>

- ・ イチローは本当に”安打製造機”かもしれない

<http://gg-hogehoge.hatenablog.com/entry/2014/01/03/143351>

目次

Rと野球データ

- ・ 打撃能力の比較1: 得点創出能力XR
- ・ 打撃能力の比較2: 野球は1人でもできる
- ・ まとめ

イチローと松井 どっちが凄い



- ・ イチロー
- ・ 高校通算打率7割
- ・ NPB 6年連続首位打者
- ・ MLB10年連続200本安打
- ・ MLB歴代シーズン最多安打
- ・ その他

イチローと松井 どっちが凄い



- ・ 松井秀喜
- ・ 高校通算60本塁打
- ・ NPB 本塁打王3回
- ・ MLB シーズン100打点4回
- ・ ワールドシリーズMVP
- ・ その他

イチローと松井 どっちが凄い

- ・ 各種データを駆使して、比べてみます

| | 投球 | 打席 | シーズン |
|---------|-----------|------------|--------|
| データ | pitch f/x | retrosheet | Lahman |
| Rのパッケージ | pitchRx | retrosheet | Lahman |

目次

- ・ Rと野球データ

打撃能力の比較1: 得点創出能力 XR

- ・ 打撃能力の比較2: 野球は1人でもできる
- ・ まとめ

打撃能力の比較 野球の指標

打撃能力の比較 野球の指標

| 代表的な指標 | 良い点 | 悪い点 |
|--------|-----------|----------|
| 打率 | 技術、走力を反映 | 単打=ホームラン |
| ホームラン数 | 技術、パワーを反映 | 単打も大事よね |
| 打点 | 勝利貢献度を反映 | チーム次第 |

打撃能力の比較 野球の指標

タイプの異なる打者を比べたい

| 代表的な指標 | 良い点 | 悪い点 |
|--------|-----------|----------|
| 打率 | 技術、走力を反映 | 単打=ホームラン |
| ホームラン数 | 技術、パワーを反映 | 単打も大事よね |
| 打点 | 勝利貢献度を反映 | チーム次第 |

打撃能力の比較 野球の指標

イチローが有利？

| 代表的な指標 | 良い点 | 悪い点 |
|--------|-----------|----------|
| 打率 | 技術、走力を反映 | 単打=ホームラン |
| ホームラン数 | 技術、パワーを反映 | 単打も大事よね |
| 打点 | 勝利貢献度を反映 | チーム次第 |

打撃能力の比較 野球の指標

松井が有利？

| 代表的な指標 | 良い点 | 悪い点 |
|--------|-----------|----------|
| 打率 | 技術、走力を反映 | 単打=ホームラン |
| ホームラン数 | 技術、パワーを反映 | 単打も大事よね |
| 打点 | 勝利貢献度を反映 | チーム次第 |

打撃能力の比較 野球の指標

松井が有利？

| 代表的な指標 | 良い点 | 悪い点 |
|--------|-----------|----------|
| 打率 | 技術、走力を反映 | 単打=ホームラン |
| ホームラン数 | 技術、パワーを反映 | 単打も大事よね |
| 打点 | 勝利貢献度を反映 | チーム次第 |

打撃能力の比較 野球の指標

“普通の”指標では物足りない

| 代表的な指標 | 良い点 | 悪い点 |
|--------|-----------|----------|
| 打率 | 技術、走力を反映 | 単打=ホームラン |
| ホームラン数 | 技術、パワーを反映 | 単打も大事よね |
| 打点 | 勝利貢献度を反映 | チーム次第 |

打撃能力の比較 野球の指標

- セイバーメトリクス
 - ・ アメリカ野球学会(SABR)の分析方法
 - ・ マネーボールで有名になった(?)
 - ・ なんだか難しそう

打撃能力の比較 野球の指標

- セイバーメトリクス 代表的な指標
 - ・ RC :
 - ・ RCWIN :
 - ・ RCAA :
 - ・ XR :

打撃能力の比較 野球の指標

- セイバーメトリクス 代表的な指標
 - ・ RC：創出得点
 - ・ RCWIN：何勝分に貢献したか
 - ・ RCAA：RC - 平均打者のRC
 - ・ XR：創出得点(改良版)

打撃能力の比較 野球の指標

- セイバーメトリクス => よく分からん
- ・ RC : 創出得点
- ・ RCWIN : 何勝分に貢献したか
- ・ RCAA : $RC - \text{平均打者のRC}$
- ・ XR : 創出得点(改良版)

打撃能力の比較 XR


- ・ XR: 創出得点

$$\begin{aligned} \text{XR} = & 0.50 \times \text{単打数} + 0.72 \times \text{二塁打数} + 1.04 \times \text{三塁打数} \\ & + 1.44 \times \text{本塁打数} + 0.34 \times \text{四死球数} + 0.25 \times \text{敬遠数} \\ & + 0.18 \times \text{盗塁数} - 0.098 \times \text{三振数} + \dots \end{aligned}$$

- ・ 各種成績の線型和

打撃能力の比較 XR

- ・ XR: 創出得点

$$\begin{aligned} \text{XR} = & 0.50 \times \text{単打数} + 0.72 \times \text{二塁打数} + 1.04 \times \text{三塁打数} \\ & + 1.44 \times \text{本塁打数} + 0.34 \times \text{四死球数} + 0.25 \times \text{敬遠数} \\ & + 0.18 \times \text{盗塁数} - 0.098 \times \text{三振数} + \dots \end{aligned}$$


打撃能力の比較 XR


- ・ XR: 創出得点

$$\begin{aligned} \text{XR} = & 0.50 \times \text{単打数} + 0.72 \times \text{二塁打数} + 1.04 \times \text{三塁打数} \\ & + 1.44 \times \text{本塁打数} + 0.34 \times \text{四死球数} + 0.25 \times \text{敬遠数} \\ & + 0.18 \times \text{盗塁数} - 0.098 \times \text{三振数} + \dots \end{aligned}$$

- ・ ホームラン1本 = 1.41点

打撃能力の比較 XR

- ・ XR: 創出得点

$$\begin{aligned} \text{XR} = & 0.50 \times \text{単打数} + 0.72 \times \text{二塁打数} + 1.04 \times \text{三塁打数} \\ & + 1.44 \times \text{本塁打数} + 0.34 \times \text{四死球数} + 0.25 \times \text{敬遠数} \\ & + 0.18 \times \text{盗塁数} - 0.098 \times \text{三振数} + \dots \end{aligned}$$


- ・ ホームラン1本 = 1.44点：価値の重み


打撃能力の比較 XR

- XR: 創出得点

| 代表的な指標 | 良い点 | 悪い点 |
|----------|-----------|----------|
| 打率 | 技術、走力を反映 | 単打=ホームラン |
| ホームラン数 | 技術、パワーを反映 | 単打も大事よね |
| 創出得点(XR) | 1人で何点取れるか | 直感的じゃない |

打撃能力の比較 XR


- ・ XR: 創出得点

$$\begin{aligned} \text{XR} = & 0.50 \times \text{単打数} + 0.72 \times \text{二塁打数} + 1.04 \times \text{三塁打数} \\ & + 1.44 \times \text{本塁打数} + 0.34 \times \text{四死球数} + 0.25 \times \text{敬遠数} \\ & + 0.18 \times \text{盗塁数} - 0.098 \times \text{三振数} + \dots \end{aligned}$$


- ・ ホームラン1本 = 1.44点：価値の重み

打撃能力の比較 XR

- ・ XR: 創出得点

$$\begin{aligned} \text{XR} = & 0.50 \times \text{単打数} + 0.72 \times \text{二塁打数} + 1.04 \times \text{三塁打数} \\ & + 1.44 \times \text{本塁打数} + 0.34 \times \text{四死球数} + 0.25 \times \text{敬遠数} \\ & + 0.18 \times \text{盗塁数} - 0.098 \times \text{三振数} + \dots \end{aligned}$$


- ・ ホームラン1本 = 1.44点 この重みの根拠は？

打撃能力の比較 XR

- ・ セイバーメトリクス XRを再現したい

1試合の得点 ~ 単打数 + 二塁打数 + 三塁打数 + 本塁打数
+ 四球数 + 三振数 + 盗塁数 + 盗塁死数 + ...

- ・ 試合結果データ + 線形回帰
重みを推定してみます

打撃能力の比較 XR

- 。利用したデータ: retrosheet (2005年)
 - ・ 約19万イベントを、97項目で記述

```
> ## devtools::install_github("gghatano/retrosheet")
> library(retrosheet)
> ## データの用意。
> dat = retrosheet(2005)
> dat %>% dim
[1] 190907      97
```

打撃能力の比較 XR

- 。利用したデータ: retrosheet (2005年)
 - ・ 約19万イベントを、97項目で記述

```
> ## devtools::install_github("gghatano/retrosheet")  
> library(retrosheet)  
> ## データの用意。  
> dat = retrosheet(2005)  
> dat %>% dim  
[1] 190907      97
```




作りました!!!!!!

打撃能力の比較 XR

- 。利用したデータ: retrosheet (2005年)
 - ・ 約19万イベントを、97項目で記述

```
> ## devtools::install_github("gghatano/retrosheet")
> library(retrosheet)
> ## データの用意。
> dat = retrosheet(2005)
> dat %>% dim
[1] 190907    97
```



2005年のデータを取得できます!!!!!!!!!!!!!!

打撃能力の比較 XR

- $Im(\text{試合の得点} \sim \text{単打数} + \text{二塁打数} + \cdots)$

打撃能力の比較 XR

- $\text{lm}(\text{試合の得点} \sim \text{単打数} + \text{二塁打数} + \dots)$

```
> lm_res = lm(dat_score_event_count, formula = SCORE ~ .)
> names(lm_coef) = nam
> lm_coef
```

| 切片 | 凡打 | 三振 | 盗塁 | 盗塁死 |
|------------|-------------|-------------|------------|-------------|
| 0.63012469 | -0.11857960 | -0.12262584 | 0.08676059 | -0.36105497 |

| 単打 | 二塁打 | 三塁打 | 本塁打 |
|-----------|-----------|-----------|-----------|
| 0.4678436 | 0.7651469 | 1.0312518 | 1.3972385 |

結果: XRと線形回帰

- XRの定義式

$$\begin{aligned} \text{XR} = & 0.50 \times \text{単打数} + 0.72 \times \text{二塁打数} + 1.04 \times \text{三塁打数} \\ & + 1.44 \times \text{本塁打数} + 0.34 \times \text{四死球数} + 0.25 \times \text{敬遠数} \\ & + 0.18 \times \text{盗塁数} - 0.098 \times \text{三振数} + \dots \end{aligned}$$

結果: XRと線形回帰

- XRの定義式

$$\begin{aligned} \text{XR} = & 0.50 \times \text{単打数} + 0.72 \times \text{二塁打数} + 1.04 \times \text{三塁打数} \\ & + 1.44 \times \text{本塁打数} + 0.34 \times \text{四死球数} + 0.25 \times \text{敬遠数} \\ & + 0.18 \times \text{盗塁数} - 0.098 \times \text{三振数} + \dots \end{aligned}$$

- 線形回帰で計算してみた結果

$$\begin{aligned} \text{XR} = & 0.47 \times \text{単打数} + 0.76 \times \text{二塁打数} + 1.03 \times \text{三塁打数} \\ & + 1.40 \times \text{本塁打数} + 0.33 \times \text{四死球数} + 0.10 \times \text{敬遠数} \\ & + 0.09 \times \text{盗塁数} - 0.12 \times \text{三振数} + \dots \end{aligned}$$

結果: XRと線形回帰

- XRの定義式

$$\begin{aligned} \text{XR} = & 0.50 \times \text{単打数} + 0.72 \times \text{二塁打数} + 1.04 \times \text{三塁打数} \\ & + 1.44 \times \text{本塁打数} + 0.34 \times \text{四死球数} + 0.25 \times \text{敬遠数} \\ & + 0.18 \times \text{盗塁数} - 0.098 \times \text{三振数} + \dots \end{aligned}$$

- 線形回帰で計算してみた結果

$$\begin{aligned} \text{XR} = & 0.47 \times \text{単打数} + 0.76 \times \text{二塁打数} + 1.03 \times \text{三塁打数} \\ & + 1.40 \times \text{本塁打数} + 0.33 \times \text{四死球数} + 0.10 \times \text{敬遠数} \\ & + 0.09 \times \text{盗塁数} - 0.12 \times \text{三振数} + \dots \end{aligned}$$

結果: XRと線形回帰

- XRの定義式

$$\begin{aligned} \text{XR} = & 0.50 \times \text{単打数} + 0.72 \times \text{二塁打数} + 1.04 \times \text{三塁打数} \\ & + 1.44 \times \text{本塁打数} + 0.34 \times \text{四死球数} + 0.25 \times \text{敬遠数} \\ & + 0.18 \times \text{盗塁数} - 0.098 \times \text{三振数} + \dots \end{aligned}$$

- 線形回帰で計算してみた結果

$$\begin{aligned} \text{XR} = & 0.47 \times \text{単打数} + 0.76 \times \text{二塁打数} + 1.03 \times \text{三塁打数} \\ & + 1.40 \times \text{本塁打数} + 0.33 \times \text{四死球数} + 0.10 \times \text{敬遠数} \\ & + 0.09 \times \text{盗塁数} - 0.12 \times \text{三振数} + \dots \end{aligned}$$

打撃能力の比較 XR

- ・ XR: 何点生み出したか
- ・ 得点 \sim 単打数 + 二塁打 + 三塁打 + \dots
- ・ 線形回帰 難しくない

http://qiita.com/gg_hatano/items/e2881eda566e21d0ba97

イチローと松井 どっちが凄い

XR: 1シーズンで何点生み出したか

| | yearID | name | HIT | HR | XR_ |
|---|--------|---------------|-----|----|---------|
| 1 | 2004 | Ichiro Suzuki | 262 | 8 | 106.106 |
| 2 | 2004 | Hideki Matsui | 174 | 31 | 103.425 |
| 3 | 2005 | Hideki Matsui | 192 | 23 | 96.433 |
| 4 | 2001 | Ichiro Suzuki | 242 | 8 | 95.228 |
| 5 | 2007 | Ichiro Suzuki | 238 | 6 | 93.975 |

イチローと松井 どっちが凄い

XR: 1シーズンで何点生み出したか

| | yearID | name | HIT | HR | XR_ |
|---|--------|---------------|-----|----|---------|
| 1 | 2004 | Ichiro Suzuki | 262 | 8 | 106.106 |
| 2 | 2004 | Hideki Matsui | 174 | 31 | 103.425 |
| 3 | 2005 | Hideki Matsui | 192 | 23 | 96.433 |
| 4 | 2001 | Ichiro Suzuki | 242 | 8 | 95.228 |
| 5 | 2007 | Ichiro Suzuki | 238 | 6 | 93.975 |

イチローと松井 どっちが凄い

XR/AB: 1打席で何点生み出したか

| | yearID | name | HIT | HR | XRperAB |
|---|--------|---------------|-----|----|---------|
| 1 | 2004 | Hideki Matsui | 174 | 31 | 0.152 |
| 2 | 2006 | Hideki Matsui | 52 | 8 | 0.147 |
| 3 | 2009 | Hideki Matsui | 125 | 28 | 0.140 |
| 4 | 2005 | Hideki Matsui | 192 | 23 | 0.137 |
| 5 | 2007 | Hideki Matsui | 156 | 25 | 0.136 |
| 6 | 2004 | Ichiro Suzuki | 262 | 8 | 0.136 |

イチローと松井 どっちが凄い

XR/AB: 1打席で何点生み出したか

| | yearID | name | | HIT | HR | XRperAB |
|---|--------|--------|--------|-----|----|---------|
| 1 | 2004 | Hideki | Matsui | 174 | 31 | 0.152 |
| 2 | 2006 | Hideki | Matsui | 52 | 8 | 0.147 |
| 3 | 2009 | Hideki | Matsui | 125 | 28 | 0.140 |
| 4 | 2005 | Hideki | Matsui | 192 | 23 | 0.137 |
| 5 | 2007 | Hideki | Matsui | 156 | 25 | 0.136 |
| 6 | 2004 | Ichiro | Suzuki | 262 | 8 | 0.136 |

イチローと松井 どっちが凄い

XR/AB: 1打席で何点生み出したか

| | yearID | name | HIT | HR | XRperAB |
|---|--------|---------------|-----|----|---------|
| 1 | 2004 | Hideki Matsui | 174 | 31 | 0.152 |
| 2 | 2006 | Hideki Matsui | 52 | 8 | 0.147 |
| 3 | 2009 | Hideki Matsui | 125 | 28 | 0.140 |
| 4 | 2005 | Hideki Matsui | 192 | 23 | 0.137 |
| 5 | 2007 | Hideki Matsui | 156 | 25 | 0.136 |
| 6 | 2004 | Ichiro Suzuki | 262 | 8 | 0.136 |

ここまでのまとめ

- 創出得点(XR): 得点重みの線型和
 - ・ 例えば、ホームラン1本=1.4点
- イチローと松井 の比較
 - ・ 通算ではイチロー、効率では松井

目次

- ・ Rと野球データ
- ・ 打撃能力の比較1: 得点創出能力XR
- ・ 打撃能力の比較2: 野球は1人でもできる
- ・ まとめ

野球のシミュレータ作った

- ・ チームスポーツなので、
個人の能力を測るのは、難しい
- ・ 1人で試合をやってもらいたい
=> 野球のシミュレータを作りました

野球のシミュレータ作った

- ・ チームスポーツなので、
個人の能力を測るのは、難しい
- ・ 1人で試合をやってもらいたい

=> シミュレータで実験しました

野球のシミュレータ作った

- ・ チームスポーツなので、
個人の能力を測るのは、難しい
- ・ 1人で試合をやってもらいたい
=> シミュレータで実験しました

野球のシミュレータ作った

- ・ 実行例 (巨人の試合)
 - ・ 村田 四球
 - ・ 長野 凡打
 - ・ アンダーソン 単打
 - ・ セペタ 三振

野球のシミュレータ作った

```
Three Out. Next Inning :: 2.  
start = "2014-04-05", end = "2014-04-05"  
-----  
Batter Name: 村田 修一  
Base: 1:0, 2: 0, 3: 0  
Result: BB  
-----  
Batter Name: 長野 久義  
Base: 1:1, 2: 0, 3: 0  
Result: OUT  
-----  
Batter Name: アンダーソン  
Base: 1:1, 2: 0, 3: 0  
Result: Single  
-----  
Batter Name: セペダ  
Base: 1:1, 2: 1, 3: 0  
Result: Swing Out
```

- ・ 実行例 (巨人の試合)
- ・ 村田 四球
- ・ 長野 凡打
- ・ アンダーソン 単打
- ・ セペタ 三振


野球のシミュレータ作った

```
Three Out. Next Inning :: 2.  
start = "2014-04-05", end = "2014-04-05"  
-----  
Batter Name: 村田 修一  
Base: 1:0, 2: 0, 3: 0  
Result: BB  
-----
```

- ・ 実行例 (巨人の試合)
- ・ 村田 四球
- ・ 長野 凡打
- ・ アンダーソン 単打
- ・ セペタ 三振

野球のシミュレータ作った

```
Three Out. Next Inning :: 2.  
start = "2014-04-05", end = "2014-04-05"  
Batter Name: 村田修一  
Base: 1:0, 2:0, 3:0  
Result: BB
```



- ・ 実行例 (巨人の試合)
- ・ 村田 四球
- ・ 長野 凡打
- ・ アンダーソン 単打
- ・ セペタ 三振

野球のシミュレータ作った

```
Three Out. Next Inning :: 2.  
start = "2014-04-05", end = "2014-04-05"  
-----  
Batter Name: 村田 修一  
Base: 1:0, 2: 0, 3: 0  
Result: BB  
-----  
Batter Name: 長野 久義  
Base: 1:1, 2: 0, 3: 0  
Result: OUT  
-----
```

- ・ 実行例 (巨人の試合)
- ・ 村田 四球
- ・ 長野 凡打
- ・ アンダーソン 単打
- ・ セペタ 三振

野球のシミュレータ作った

```
Three Out. Next Inning :: 2.  
start = "2014-04-05", end = "2014-04-05"  
-----  
Batter Name: 村田 修一  
Base: 1:0, 2: 0, 3: 0  
Result: BB  
-----  
Batter Name: 長野 久義  
Base: 1:1, 2: 0, 3: 0  
Result: OUT  
-----
```

- ・ 実行例 (巨人の試合)

- ・ 村田 四球


- ・ 長野 凡打

- ・ アンダーソン 単打

- ・ セペタ 三振

野球のシミュレータ作った

```
Three Out. Next Inning :: 2.  
start = "2014-04-05", end = "2014-04-05"  
-----  
Batter Name: 村田 修一  
Base: 1:0, 2: 0, 3: 0  
Result: BB  
-----  
Batter Name: 長野 久義  
Base: 1:1, 2: 0, 3: 0  
Result: OOI  
-----
```



- ・ 実行例 (巨人の試合)

- ・ 村田 四球

- ・ 長野 凡打

- ・ アンダーソン 単打

- ・ セペタ 三振

野球のシミュレータ作った

```
Three Out. Next Inning :: 2.  
start = "2014-04-05", end = "2014-04-05"  
-----  
Batter Name: 村田 修一  
Base: 1:0, 2: 0, 3: 0  
Result: BB  
-----  
Batter Name: 長野 久義  
Base: 1:1, 2: 0, 3: 0  
Result: OUT  
-----  
Batter Name: アンダーソン  
Base: 1:1, 2: 0, 3: 0  
Result: Single  
-----
```

- ・ 実行例 (巨人の試合)
- ・ 村田 四球
- ・ 長野 凡打
- ・ アンダーソン 単打
- ・ セペタ 三振

野球のシミュレータ作った

```
Three Out. Next Inning :: 2.  
start = "2014-04-05", end = "2014-04-05"  
-----  
Batter Name: 村田 修一  
Base: 1:0, 2: 0, 3: 0  
Result: BB  
-----  
Batter Name: 長野 久義  
Base: 1:1, 2: 0, 3: 0  
Result: OUT  
-----  
Batter Name: アンダーソン  
Base: 1:1, 2: 0, 3: 0  
Result: Single  
-----  
Batter Name: セペダ  
Base: 1:1, 2: 1, 3: 0  
Result: Swing Out
```

- ・ 実行例 (巨人の試合)
- ・ 村田 四球
- ・ 長野 凡打
- ・ アンダーソン 単打
- ・ セペダ 三振

野球のシミュレータ作った

```
Three Out. Next Inning :: 2.  
start = "2014-04-05", end = "2014-04-05"  
-----  
Batter Name: 村田 修一  
Base: 1:0, 2: 0, 3: 0  
Result: BB  
-----  
Batter Name: 長野 久義  
Base: 1:1, 2: 0, 3: 0  
Result: OUT  
-----  
Batter Name: アンダーソン  
Base: 1:1, 2: 0, 3: 0  
Result: Single  
-----  
Batter Name: セペダ  
Base: 1:1, 2: 1, 3: 0  
Result: Swing Out
```

- ・ 実行例 (巨人の試合)
- ・ 村田 四球
- ・ 長野 凡打
- ・ アンダーソン 単打
- ・ セペダ 三振
- ・ チェンジ

野球のシミュレータ作った

```
Three Out. Next Inning :: 2.  
start = "2014-04-05", end = "2014-04-05"  
-----  
Batter Name: 村田 修一  
Base: 1:0, 2: 0, 3: 0  
Result: BB  
-----  
Batter Name: 長野 久義  
Base: 1:1, 2: 0, 3: 0  
Result: OUT  
-----  
Batter Name: アンダーソン  
Base: 1:1, 2: 0, 3: 0  
Result: Single  
-----  
Batter Name: セペダ  
Base: 1:1, 2: 1, 3: 0  
Result: Swing Out
```

- ・ 実行例 (巨人の試合)
- ・ 村田 四球
- ・ 長野 凡打
- ・ アンダーソン 単打
- ・ セペダ 三振
- ・ チェンジ(怒)

野球のシミュレータ作った

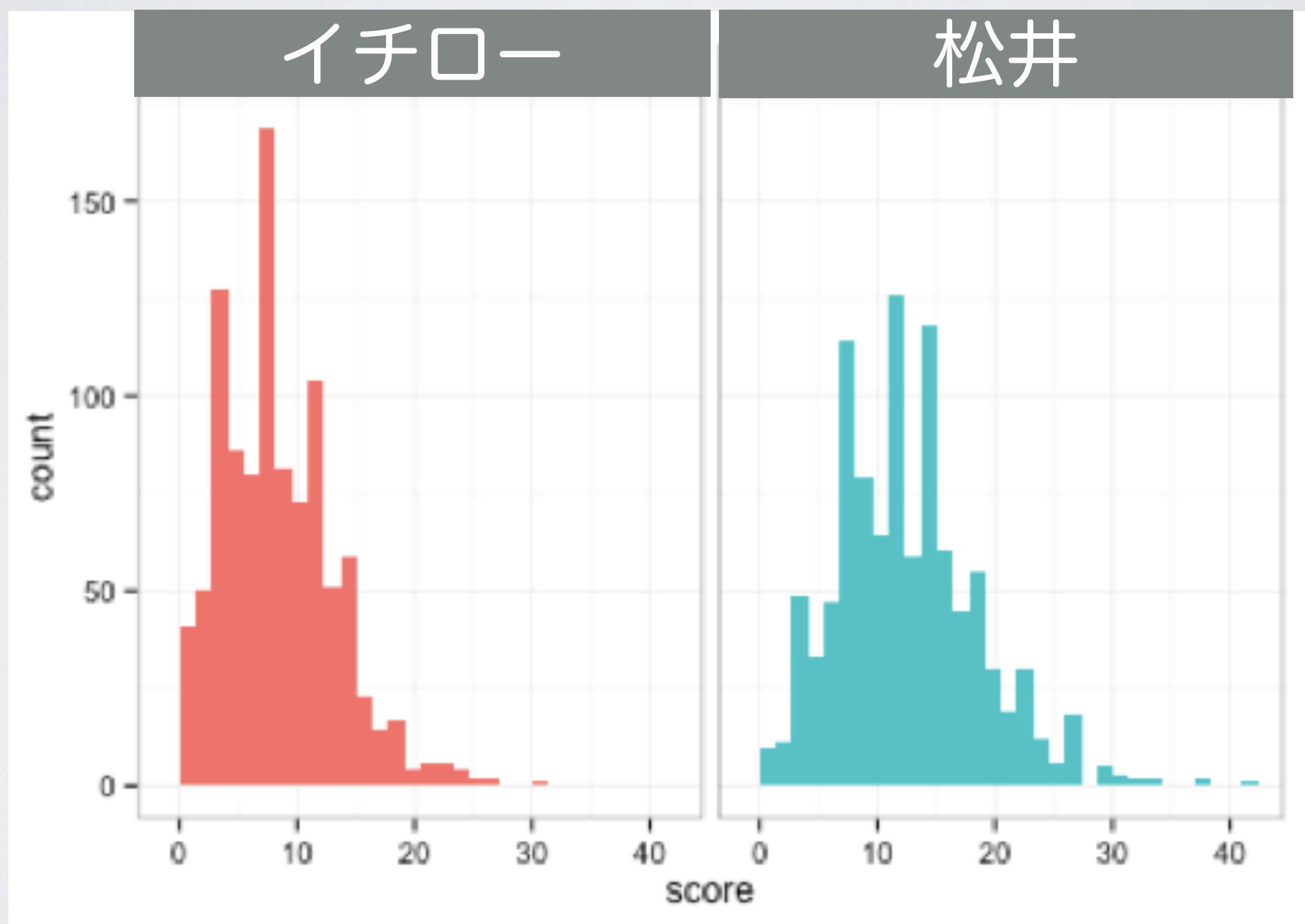
1. メンバーを入力
2. 選手の個人成績を取得
3. 個人成績と乱数で、打撃結果を決定
4. 1試合やって、何点取れたかを見る

1000試合ずつ実行します



1試合での得点 分布

1試合での得点 分布



1試合での得点 平均

イチロー: 8.4点, 松井: 12.6点

```
> dat_result %>% group_by(name) %>%  
+ summarise(score = mean(score)) %>%  
+ as.data.frame  
  name  score  
1 Ichiro  8.406  
2 Matsui 12.612
```


1試合での得点 平均

イチロー: 8.4点, 松井: 12.6点

```
> dat_result %>% group_by(name) %>%  
+ summarise(score = mean(score)) %>%  
+ as.data.frame
```

| | name | score |
|---|--------|--------|
| 1 | Ichiro | 8.406 |
| 2 | Matsui | 12.612 |

1試合での得点 平均

XRで計算される得点数と比べてみる

```
> dat_result %>% group_by(name) %>%  
+ summarise(score = mean(score),  
+ score_XR = mean(score_XR)) %>%  
+ as.data.frame
```

| | name | score | score_XR |
|---|--------|--------|----------|
| 1 | Ichiro | 8.406 | 6.554892 |
| 2 | Matsui | 12.612 | 7.842793 |

1試合での得点 平均

XRから計算される得点とはズレている

```
> dat_result %>% group_by(name) %>%  
+ summarise(score = mean(score),  
+ score_XR = mean(score_XR)) %>%  
+ as.data.frame
```

| | name | score | score_XR |
|---|--------|--------|----------|
| 1 | Ichiro | 8.406 | 6.554892 |
| 2 | Matsui | 12.612 | 7.842793 |

目次

- ・ Rと野球データ
- ・ 打撃能力の比較1: 得点創出能力XR
- ・ 打撃能力の比較2: 野球は1人でもできる

まとめ

発表の目的 (再掲)

1. セイバーメトリクスに親しみたい(僕が)
2. Rで野球データを触れることを伝えたい
3. 作ったパッケージの宣伝をしたい

まとめ

- ・ Rで野球のデータ取得、分析が出来ました
- ・ セイバーメトリクスの実験もできました
- ・ 今回の実験では、

通算 イチロー > 松井, 効率 イチロー < 松井

まとめ

- ・ Rで野球のデータ取得、分析が出来ました
- ・ セイバーメトリクスの実験もできました
- ・ 今回の実験では、
通算 イチロー > 松井, 効率 イチロー < 松井
- ・ ありがとうございました。