Using online update algorithms to predict speedway results.

Dawid Kałędkowski @ClickMeeting

Speedway is a motorcycle sport.

- **Motorbike** racing on **dirt** circuit.

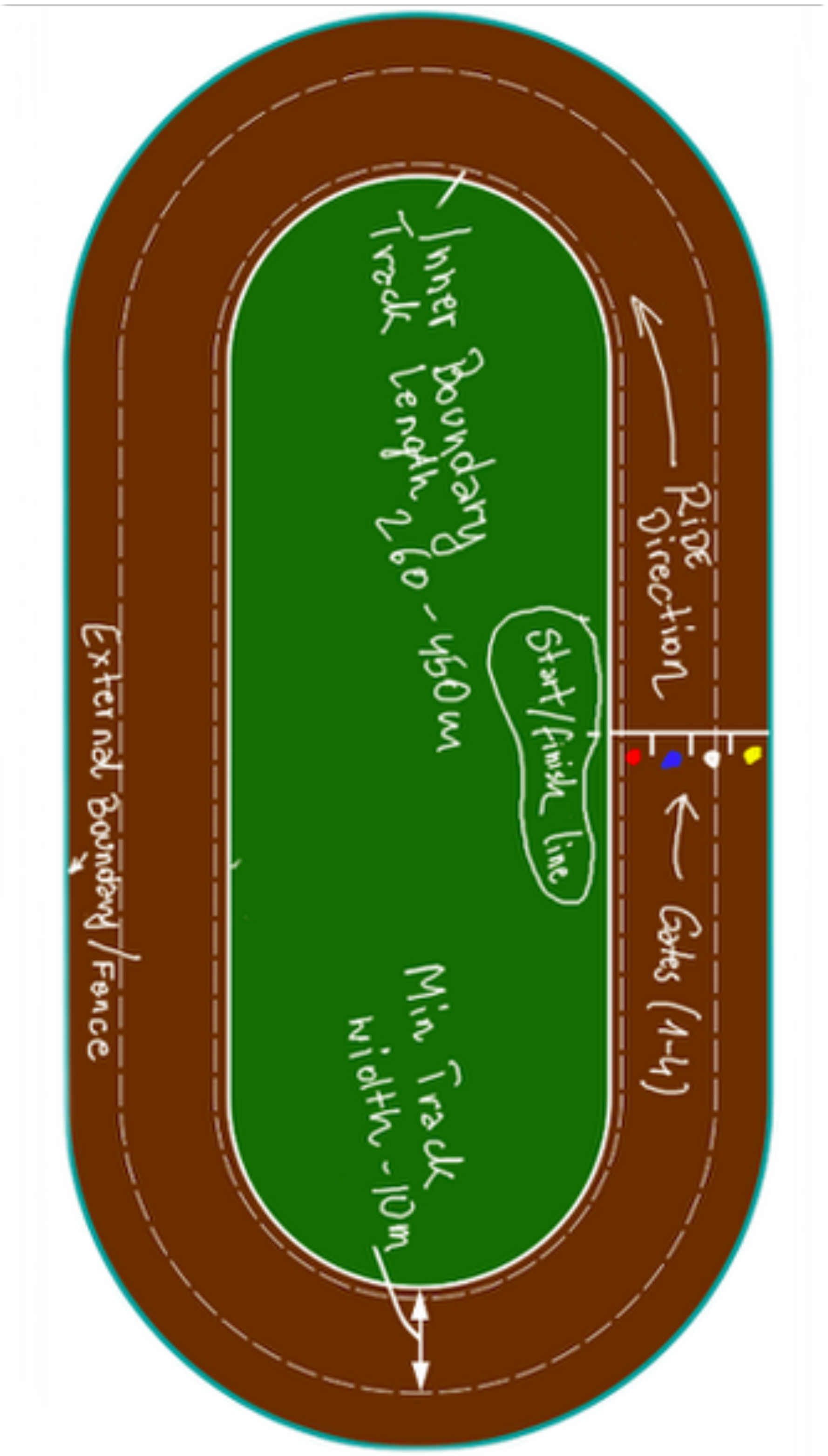
- No gears and **no breaks**.

- 4 riders competing over four laps.

- Start indicated by **tape rise** and light signal.

- Exclusions for falling, engine failure, retiring, touching tape etc.

- Sliding scale for scoring (known as the **3-2-1-0 method**)



More than 1000 events annualy.

- 1000 riders from 31 countries

- 100 different competitions.

- Individual competitions - Grand-Prix

- League matches - PGE Ekstraliga, Premiership, Elitserien, Bundesliga etc.

- Team competitions - SWC/SON

- Pair competitions - SBP

Statistics matter.

Grand Prix 2006

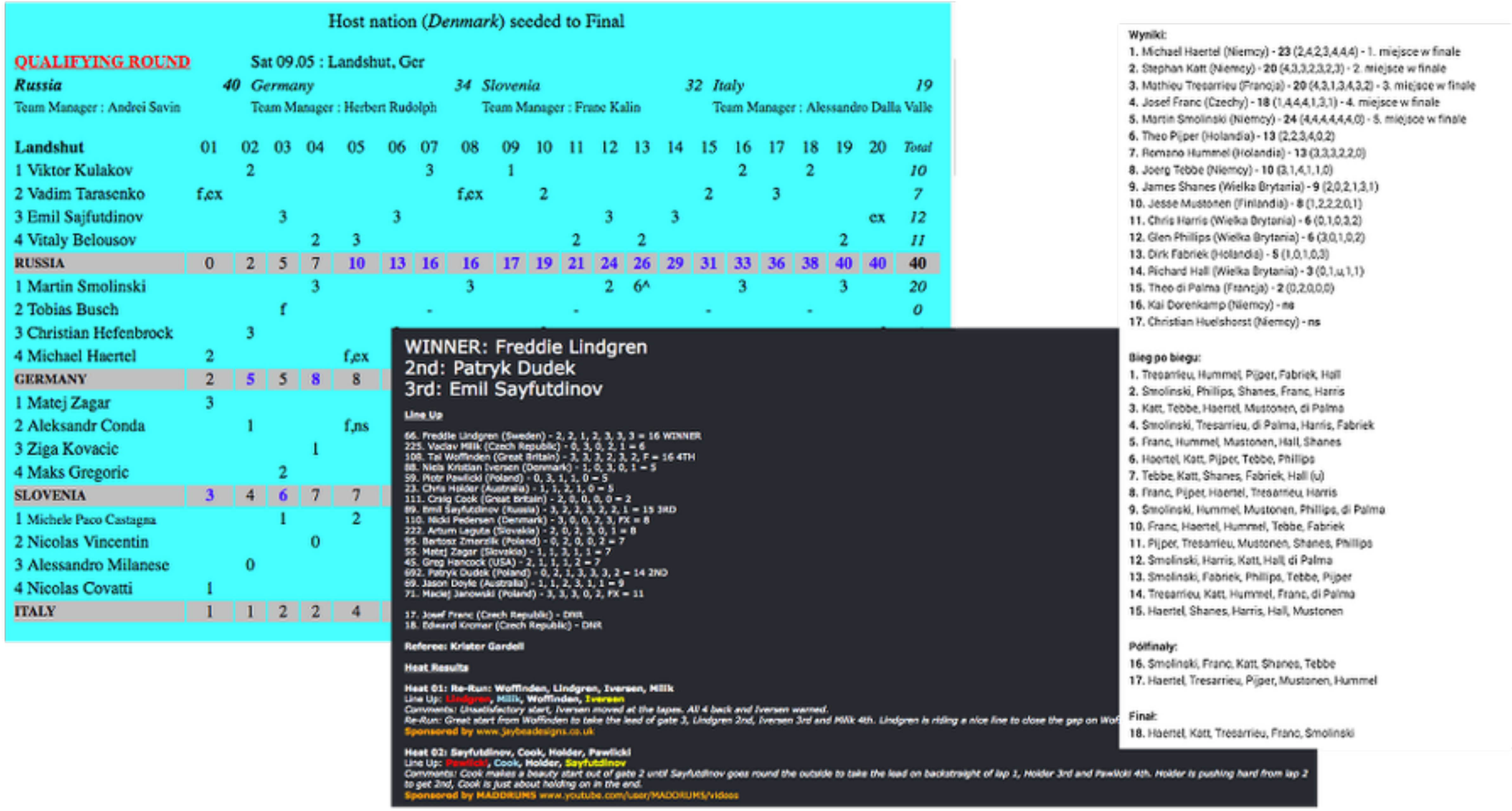
Bieg	Tor	Nr	ZAWODNIK	Miejsce	Punkty
1	🟥	2	Jason CRUMP	[illegible]	[illegible]
	🟦	9	Scott NICHOLLS	[illegible]	[illegible]
	⬜	3	Leigh ADAMS	[illegible]	[illegible]
	🟨	15	Niels Kristian IVERSEN	[illegible]	[illegible]
2	🟥	13	Matej ZAGAR	[illegible]	[illegible]
	🟦	8	Andreas JONSSON	[illegible]	[illegible]
	⬜	14	Piotr PROTASIEWICZ	[illegible]	[illegible]
	🟨	12	Lee RICHARDSON	[illegible]	[illegible]
3	🟥	19	Hans ANDERSEN	[illegible]	[illegible]
	🟦	11	Jarosław HAMPEL	[illegible]	[illegible]
	⬜	7	Tomasz GOLLOB	[illegible]	[illegible]
	🟨	16	Wiesław JAGUŚ	[illegible]	[illegible]
4	🟥	6	Bjarne PEDERSEN	[illegible]	[illegible]
	🟦	4	Nicki PEDERSEN	[illegible]	[illegible]
	⬜	10	Antonio LINDBAECK	[illegible]	[illegible]
	🟨	5	Greg HANCOCK	[illegible]	[illegible]
5	🟥	5	Greg HANCOCK	[illegible]	[illegible]
	🟦	2	Jason CRUMP	[illegible]	[illegible]
	⬜	13	Matej ZAGAR	[illegible]	[illegible]
	🟨	7	Tomasz GOLLOB	[illegible]	[illegible]
6	🟥	4	Nicki PEDERSEN	[illegible]	[illegible]
	🟦	19	Hans ANDERSEN	[illegible]	[illegible]
	⬜	9	Scott NICHOLLS	[illegible]	[illegible]
	🟨	14	Piotr PROTASIEWICZ	[illegible]	[illegible]
7	🟥	11	Jarosław HAMPEL	[illegible]	[illegible]
	🟦	6	Bjarne PEDERSEN	[illegible]	[illegible]
	⬜	8	Andreas JONSSON	[illegible]	[illegible]
	🟨	3	Leigh ADAMS	[illegible]	[illegible]
8	🟥	15	Niels Kristian IVERSEN	[illegible]	[illegible]
	🟦	12	Lee RICHARDSON	[illegible]	[illegible]
	⬜	16	Wiesław JAGUŚ	[illegible]	[illegible]
	🟨	10	Antonio LINDBAECK	[illegible]	[illegible]
9	🟥	14	Piotr PROTASIEWICZ	[illegible]	[illegible]
	🟦	10	Antonio LINDBAECK	[illegible]	[illegible]
	⬜	2	Jason CRUMP	[illegible]	[illegible]
	🟨	11	Jarosław HAMPEL	[illegible]	[illegible]
10	🟥	16	Wiesław JAGUŚ	[illegible]	[illegible]
	🟦	13	Matej ZAGAR	[illegible]	[illegible]
	⬜	6	Bjarne PEDERSEN	[illegible]	[illegible]
	🟨	9	Scott NICHOLLS	[illegible]	[illegible]
11	🟥	12	Lee RICHARDSON	[illegible]	[illegible]
	🟦	7	Tomasz GOLLOB	[illegible]	[illegible]
	⬜	3	Leigh ADAMS	[illegible]	[illegible]
	🟨	4	Nicki PEDERSEN	[illegible]	[illegible]
12	🟥	5	Greg HANCOCK	[illegible]	[illegible]
	🟦	15	Niels Kristian IVERSEN	[illegible]	[illegible]
	⬜	19	Hans ANDERSEN	[illegible]	[illegible]
	🟨	8	Andreas JONSSON	[illegible]	[illegible]
13	🟥	8	Andreas JONSSON	[illegible]	[illegible]
	🟦	16	Wiesław JAGUŚ	[illegible]	[illegible]
	⬜	4	Nicki PEDERSEN	[illegible]	[illegible]
	🟨	2	Jason CRUMP	[illegible]	[illegible]
14	🟥	9	Scott NICHOLLS	[illegible]	[illegible]
	🟦	5	Greg HANCOCK	[illegible]	[illegible]
	⬜	12	Lee RICHARDSON	[illegible]	[illegible]
	🟨	11	Jarosław HAMPEL	[illegible]	[illegible]
15	🟥	10	Antonio LINDBAECK	[illegible]	[illegible]
	🟦	3	Leigh ADAMS	[illegible]	[illegible]
	⬜	19	Hans ANDERSEN	[illegible]	[illegible]
	🟨	13	Matej ZAGAR	[illegible]	[illegible]
16	🟥	7	Tomasz GOLLOB	[illegible]	[illegible]
	🟦	14	Piotr PROTASIEWICZ	[illegible]	[illegible]
	⬜	15	Niels Kristian IVERSEN	[illegible]	[illegible]
	🟨	6	Bjarne PEDERSEN	[illegible]	[illegible]
17	🟥	2	Jason CRUMP	[illegible]	[illegible]
	🟦	12	Lee RICHARDSON	[illegible]	[illegible]
	⬜	6	Bjarne PEDERSEN	[illegible]	[illegible]
	🟨	19	Hans ANDERSEN	[illegible]	[illegible]
18	🟥	7	Tomasz GOLLOB	[illegible]	[illegible]
	🟦	9	Scott NICHOLLS	[illegible]	[illegible]
	⬜	8	Andreas JONSSON	[illegible]	[illegible]
	🟨	10	Antonio LINDBAECK	[illegible]	[illegible]
19	🟥	3	Leigh ADAMS	[illegible]	[illegible]
	🟦	16	Wiesław JAGUŚ	[illegible]	[illegible]
	⬜	5	Greg HANCOCK	[illegible]	[illegible]
	🟨	14	Piotr PROTASIEWICZ	[illegible]	[illegible]
20	🟥	13	Matej ZAGAR	[illegible]	[illegible]
	🟦	4	Nicki PEDERSEN	[illegible]	[illegible]
	⬜	11	Jarosław HAMPEL	[illegible]	[illegible]
	🟨	15	Niels Kristian IVERSEN	[illegible]	[illegible]
21	🟥		Pedersen	[illegible]	[illegible]
	🟦		[illegible]	[illegible]	[illegible]
	⬜		Jonsson	[illegible]	[illegible]
	🟨		Gollob	[illegible]	[illegible]
22	🟥		Hancock	[illegible]	[illegible]
	🟦		[illegible]	[illegible]	[illegible]
	⬜		Crump	[illegible]	[illegible]
	🟨		Jaguś	[illegible]	[illegible]
23	🟥		Pedersen	[illegible]	[illegible]
	🟦		Gollob	[illegible]	[illegible]
	⬜		Jaguś	[illegible]	[illegible]
	🟨		Hancock	[illegible]	[illegible]

Nr	ZAWODNIK	KRAJ	1	2	3	4	5	[illegible]	[illegible]	[illegible]	[illegible]
2	Jason CRUMP	Australia	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	181	7	[illegible]
3	Leigh ADAMS	Australia	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	6	100	6	[illegible]
4	Nicki PEDERSEN	Dania	3	5	[illegible]	[illegible]	[illegible]	[illegible]	109	[illegible]	[illegible]
5	Greg HANCOCK	USA	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	124	[illegible]	[illegible]
6	Bjarne PEDERSEN	Dania	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	75	[illegible]	[illegible]
7	Tomasz GOLLOB	Polska	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	76	[illegible]	[illegible]
8	Andreas JONSSON	Szwecja	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	107	[illegible]	[illegible]
9	Scott NICHOLLS	Anglia	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	76	[illegible]	[illegible]
10	Antonio LINDBAECK	Szwecja	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	86	[illegible]	[illegible]
11	Jarosław HAMPEL	Polska	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	85	6	[illegible]
12	Lee RICHARDSON	Anglia	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	38	[illegible]	[illegible]
13	Matej ZAGAR	Słowenia	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	89	[illegible]	[illegible]
14	Piotr PROTASIEWICZ	Polska	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	27	[illegible]	[illegible]
15	Niels K. IVERSEN	Dania	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	48	[illegible]	[illegible]
16	Wiesław JAGUŚ	Polska	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-	[illegible]	[illegible]
19	Hans ANDERSEN	Dania	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	82	[illegible]	[illegible]

Zawodnicy rezerwowi

17	Janusz KOŁODZIEJ	Polska						[illegible]			
18	Karol ZĄBIK	Polska						[illegible]			

Approach to data is old-fashioned



Host nation (*Denmark*) seeded to Final

QUALIFYING ROUND — Sat 09.05 : Landshut, Ger

Russia	**40**	*Germany*	34	*Slovenia*	32	*Italy* 19
Team Manager : Andrei Savin		Team Manager : Herbert Rudolph		Team Manager : Franc Kalin		Team Manager : Alessandro Dalla Valle

Landshut	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	*Total*
1 Viktor Kulakov		2					3		1							2		2			*10*
2 Vadim Tarasenko	f,ex							f,ex		2					2		3				*7*
3 Emil Sajfutdinov			3			3						3		3						ex	*12*
4 Vitaly Belousov				2	3						2		2						2		*11*
RUSSIA	0	2	5	7	**10**	**13**	**16**	**16**	**17**	**19**	**21**	**24**	**26**	**29**	**31**	**33**	**36**	**38**	**40**	**40**	**40**
1 Martin Smolinski				3				3				2	6^			3			3		*20*
2 Tobias Busch			f				-				-				-			-			*0*
3 Christian Hefenbrock		3																			
4 Michael Haertel	2				f,ex																
GERMANY	2	**5**	5	**8**	8																
1 Matej Zagar	3																				
2 Aleksandr Conda		1			f,ns																
3 Ziga Kovacic				1																	
4 Maks Gregoric			2																		
SLOVENIA	**3**	4	**6**	7	7																
1 Michele Paco Castagna			1		2																
2 Nicolas Vincentin				0																	
3 Alessandro Milanese		0																			
4 Nicolas Covatti	1																				
ITALY	1	1	2	2	4																

WINNER: Freddie Lindgren
2nd: Patryk Dudek
3rd: Emil Sayfutdinov

Line Up

66. Freddie Lindgren (Sweden) - 2, 2, 1, 2, 3, 3, 3 = 16 WINNER
225. Vaclav Milik (Czech Republic) - 0, 3, 0, 2, 1 = 6
108. Tai Woffinden (Great Britain) - 3, 3, 3, 2, 3, 2, F = 16 4TH
88. Niels Kristian Iversen (Denmark) - 1, 0, 3, 0, 1 = 5
59. Piotr Pawlicki (Poland) - 0, 3, 1, 1, 0 = 5
23. Chris Holder (Australia) - 1, 1, 2, 1, 0 = 5
111. Craig Cook (Great Britain) - 2, 0, 0, 0, 0 = 2
89. Emil Sayfutdinov (Russia) - 3, 2, 2, 3, 2, 2, 1 = 15 3RD
110. Nicki Pedersen (Denmark) - 3, 0, 0, 2, 3, FX = 8
222. Artum Laguta (Slovakia) - 2, 0, 2, 3, 0, 1 = 8
95. Bartosz Zmarzlik (Poland) - 0, 2, 0, 0, 2 = 7
55. Matej Zagar (Slovakia) - 1, 1, 3, 1, 1 = 7
45. Greg Hancock (USA) - 2, 1, 1, 1, 2 = 7
692. Patryk Dudek (Poland) - 0, 2, 1, 3, 3, 3, 2 = 14 2ND
69. Jason Doyle (Australia) - 1, 1, 2, 3, 1, 1 = 9
71. Maciej Janowski (Poland) - 3, 3, 3, 0, 2, FX = 11

17. Josef Franc (Czech Republic) - DNR
18. Edward Kromer (Czech Republic) - DNR

Referee: Krister Gardell

Heat Results

Heat 01: Re-Run: Woffinden, Lindgren, Iversen, Milik
Line Up: **Lindgren**, **Milik**, **Woffinden**, **Iversen**
Comments: Unsatisfactory start, Iversen moved at the tapes. All 4 back and Iversen warned.
Re-Run: Great start from Woffinden to take the lead of gate 3, Lindgren 2nd, Iversen 3rd and Milik 4th. Lindgren is riding a nice line to close the gap on Wo
Sponsored by www.jaybeadesigns.co.uk

Heat 02: Sayfutdinov, Cook, Holder, Pawlicki
Line Up: **Pawlicki**, **Cook**, **Holder**, **Sayfutdinov**
Comments: Cook makes a beauty start out of gate 2 until Sayfutdinov goes round the outside to take the lead on backstraight of lap 1, Holder 3rd and Pawlicki 4th. Holder is pushing hard from lap 2 to get 2nd, Cook is just about holding on in the end.
Sponsored by MADDRUMS www.youtube.com/user/MADDRUMS/videos

Wyniki:

1. Michael Haertel (Niemcy) - **23** (2,4,2,3,4,4,4) - 1. miejsce w finale
2. Stephan Katt (Niemcy) - **20** (4,3,3,2,3,2,3) - 2. miejsce w finale
3. Mathieu Tresarrieu (Francja) - **20** (4,3,1,3,4,3,2) - 3. miejsce w finale
4. Josef Franc (Czechy) - **18** (1,4,4,4,1,3,1) - 4. miejsce w finale
5. Martin Smolinski (Niemcy) - **24** (4,4,4,4,4,4,0) - 5. miejsce w finale
6. Theo Pijper (Holandia) - **13** (2,2,3,4,0,2)
7. Romano Hummel (Holandia) - **13** (3,3,3,2,2,0)
8. Joerg Tebbe (Niemcy) - **10** (3,1,4,1,1,0)
9. James Shanes (Wielka Brytania) - **9** (2,0,2,1,3,1)
10. Jesse Mustonen (Finlandia) - **8** (1,2,2,2,0,1)
11. Chris Harris (Wielka Brytania) - **6** (0,1,0,3,2)
12. Glen Phillips (Wielka Brytania) - **6** (3,0,1,0,2)
13. Dirk Fabriek (Holandia) - **5** (1,0,1,0,3)
14. Richard Hall (Wielka Brytania) - **3** (0,1,u,1,1)
15. Theo di Palma (Francja) - **2** (0,2,0,0,0)
16. Kai Dorenkamp (Niemcy) - **ns**
17. Christian Huelshorst (Niemcy) - **ns**

Bieg po biegu:

1. Tresarrieu, Hummel, Pijper, Fabriek, Hall
2. Smolinski, Phillips, Shanes, Franc, Harris
3. Katt, Tebbe, Haertel, Mustonen, di Palma
4. Smolinski, Tresarrieu, di Palma, Harris, Fabriek
5. Franc, Hummel, Mustonen, Hall, Shanes
6. Haertel, Katt, Pijper, Tebbe, Phillips
7. Tebbe, Katt, Shanes, Fabriek, Hall (u)
8. Franc, Pijper, Haertel, Tresarrieu, Harris
9. Smolinski, Hummel, Mustonen, Phillips, di Palma
10. Franc, Haertel, Hummel, Tebbe, Fabriek
11. Pijper, Tresarrieu, Mustonen, Shanes, Phillips
12. Smolinski, Harris, Katt, Hall, di Palma
13. Smolinski, Fabriek, Phillips, Tebbe, Pijper
14. Tresarrieu, Katt, Hummel, Franc, di Palma
15. Haertel, Shanes, Harris, Hall, Mustonen

Półfinały:

16. Smolinski, Franc, Katt, Shanes, Tebbe
17. Haertel, Tresarrieu, Pijper, Mustonen, Hummel

Finał:

18. Haertel, Katt, Tresarrieu, Franc, Smolinski

source: sportowefakty.pl, speedwayresults.com, speedwayupdates.proboards.com

Data gathering partially automated.

- Multiple scrapers for different pages (`rvest`, `RSelenium` and a lot of REGEXP)

- PDF parsers (`tabulizer`)

- Finding appropriate rider-heat pattern

- Additional verifications

- Approximate string matching for language specific letters (`stringdist:amatch`)

- Apps to input and browse data (`shiny`,`rhandsontable`,`RMySQL`)

Actual speedway database

- 10528 events in 148 competitions.

- Almost 700k individual performances.

- Unified rider names, competitions and places.

- 241 Speedway stadiums with coordinates

Analytical challange

- Output is a **ranking**.

- Need for **continuous updates**.

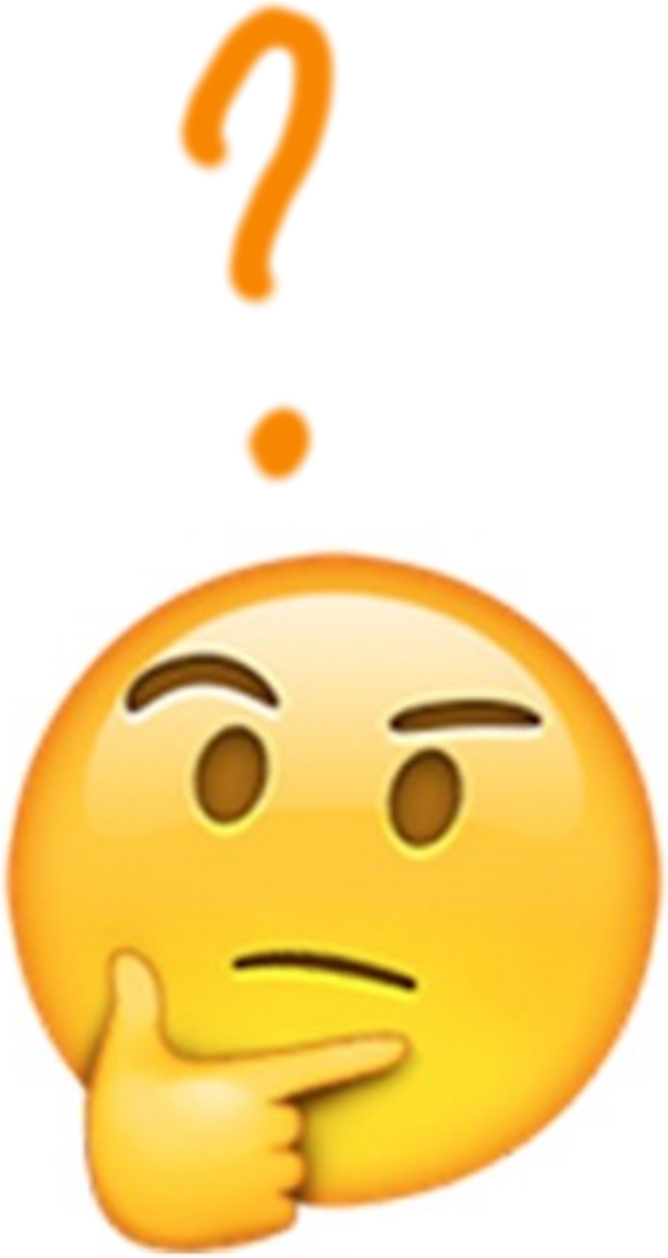
- Some **commonly known effects** need to be examined:

  - Riders form changes in time.

  - Many interactions (gate*heat*stadium).

  - Field advantage.

  - Winter break



Online update algorithms.

- Ranking modelled as **Rank Ordered Logit/BT Model.**

- Models are estimated using **Bayesian Approximation Method.**

$$R_i' \leftarrow R_i + K * (Y_i - \hat{Y}_i)$$

- Update doesn't require previous data.

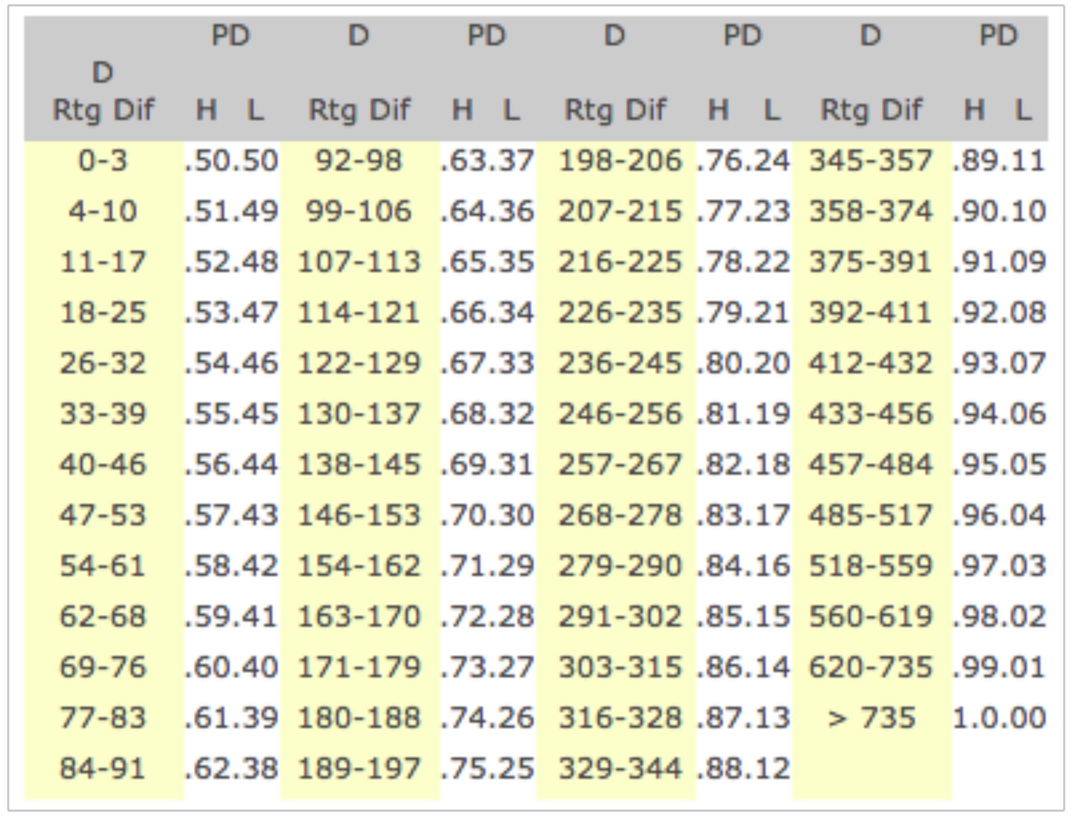
- Computationaly efficient.

Elo rating system

- Rating calculated using formula specified by FIDE.

- Points gain depends on rating difference

- No rating deviation in Elo formula



D	PD	D	PD	D	PD	D	PD
Rtg Dif	H L	Rtg Dif	H L	Rtg Dif	H L	Rtg Dif	H L
0-3	.50.50	92-98	.63.37	198-206	.76.24	345-357	.89.11
4-10	.51.49	99-106	.64.36	207-215	.77.23	358-374	.90.10
11-17	.52.48	107-113	.65.35	216-225	.78.22	375-391	.91.09
18-25	.53.47	114-121	.66.34	226-235	.79.21	392-411	.92.08
26-32	.54.46	122-129	.67.33	236-245	.80.20	412-432	.93.07
33-39	.55.45	130-137	.68.32	246-256	.81.19	433-456	.94.06
40-46	.56.44	138-145	.69.31	257-267	.82.18	457-484	.95.05
47-53	.57.43	146-153	.70.30	268-278	.83.17	485-517	.96.04
54-61	.58.42	154-162	.71.29	279-290	.84.16	518-559	.97.03
62-68	.59.41	163-170	.72.28	291-302	.85.15	560-619	.98.02
69-76	.60.40	171-179	.73.27	303-315	.86.14	620-735	.99.01
77-83	.61.39	180-188	.74.26	316-328	.87.13	> 735	1.0.00
84-91	.62.38	189-197	.75.25	329-344	.88.12		

source: https://www.fide.com/fide/handbook.html?id=172&view=article

Glicko rating system (`sport::glicko`)

- First bayesian rating system.

- Rating change depends on ratings (R) and ratings deviation (RD).

$$\hat{Y}_i = P(X_i > X_q) = \frac{1}{1 + 10^{-g(RD_{iq})*(R_i - R_q)/400}}$$

$$R'_i = R_i + \frac{1}{\frac{1}{RD_i^2} + \frac{1}{d_i^2}} * \sum_j g(RD_j) * (Y_{ij} - \hat{Y}_{ij})$$

$$RD'_i = \sqrt{(\frac{1}{RD_i^2} + \frac{1}{d_i^2})^{-1}}$$

Glicko2 rating system *(`sport::glicko2`)*

- Volatile parameter σ added. Measures expected fluctuations.

- Updated rating deviation based on the 'Illinois Algorithm'.

$$\hat{Y}_{ij} = \frac{1}{1 + e^{-g(\phi_{ij})*(\mu_i-\mu_j)}}$$

$$\phi'_i = \frac{1}{\sqrt{\frac{1}{{\phi_i}^2+{\sigma'_i}^2} + \frac{1}{v}}}$$

$$\mu'_i = \mu_i + \phi'_i * \sum_j g(\phi_j) * (Y_{ij} - \hat{Y}_{ij})$$

Bayesian Bradley Terry (`sport::bbt`)

- Extends algorithms to multi-player teams.

- Team rating/variance is a sum of team players ratings/variances.

$$\hat{Y}_{ij} = P(X_i > X_j) = \frac{e^{R_i/c_{i_j}}}{e^{R_i/c_{ij}} + e^{R_j/c_{ij}}}$$

$$R'_i = R_i + \sum_j \frac{RD_i^2}{c_{ij}} * (Y_{ij} - \hat{Y}_{ij})$$

$$RD'_i = RD_i * [1 - \frac{RD_{ij}^2}{RD_i^2} \sum_j \gamma_j * (\frac{RD_i}{c_{ij}})^2 * \hat{Y}_{ij}\hat{Y}_{ji}]$$

Bayesian Dynamic Logit (`sport::bdl`)

- EKF with logistic function as measurement equation.

- Teams/players are treated as alternative in discrete choice models.

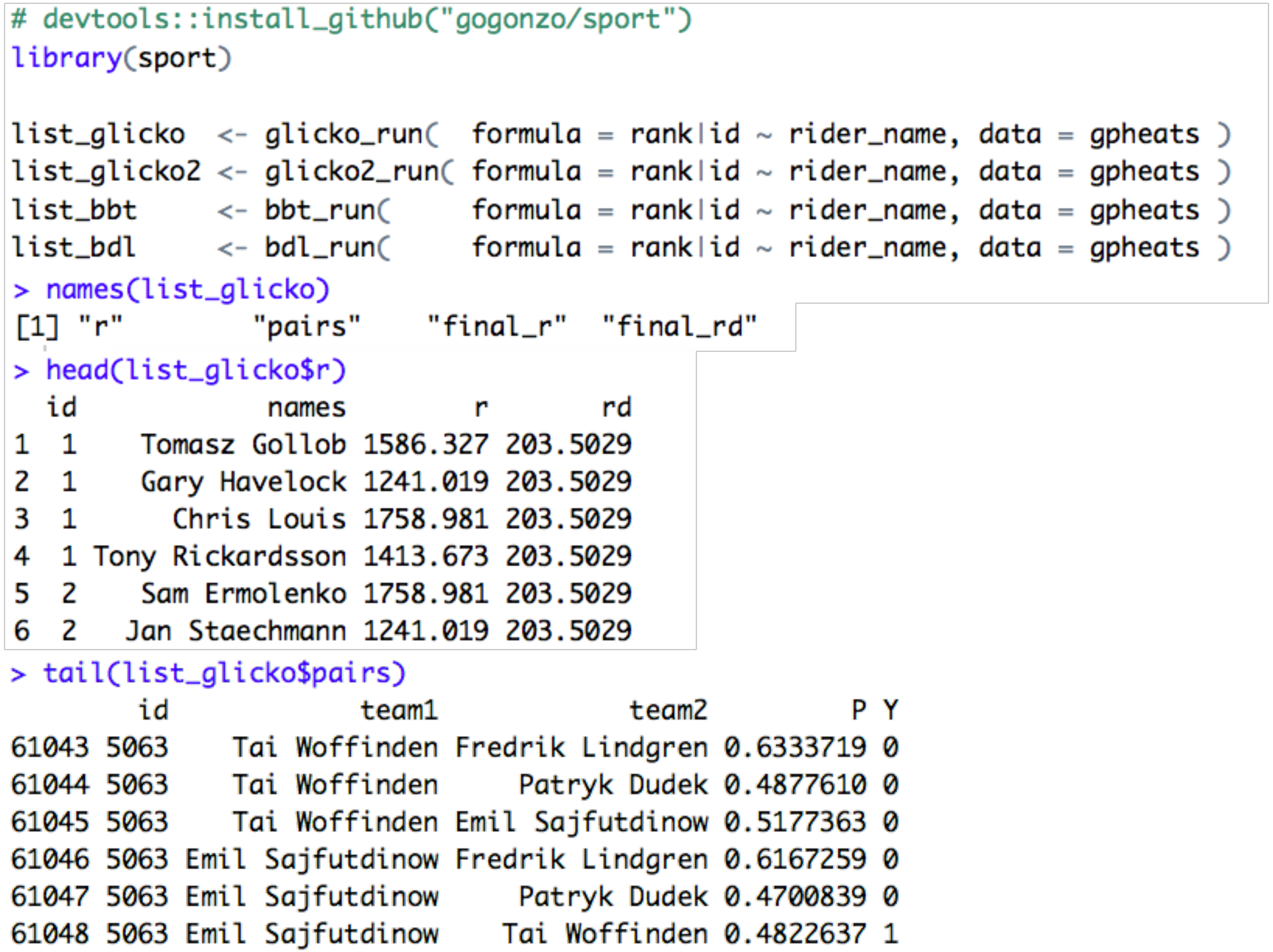
$$Y_t = \frac{e^{z}}{1 + e^{z_t}}$$

$$z_t = \beta_{it}^T x_{it} - \beta_{jt}^T x_{jt}$$

$$\hat{\boldsymbol{w}}_t = \hat{\boldsymbol{w}}_{t-1} + \boldsymbol{\Sigma}_t \boldsymbol{x}_t (z_t - y_t)$$

$$s_t^2 = \boldsymbol{x}_t^T (\boldsymbol{\Sigma}_{t-1} + q_t \boldsymbol{I}) \boldsymbol{x}_t$$

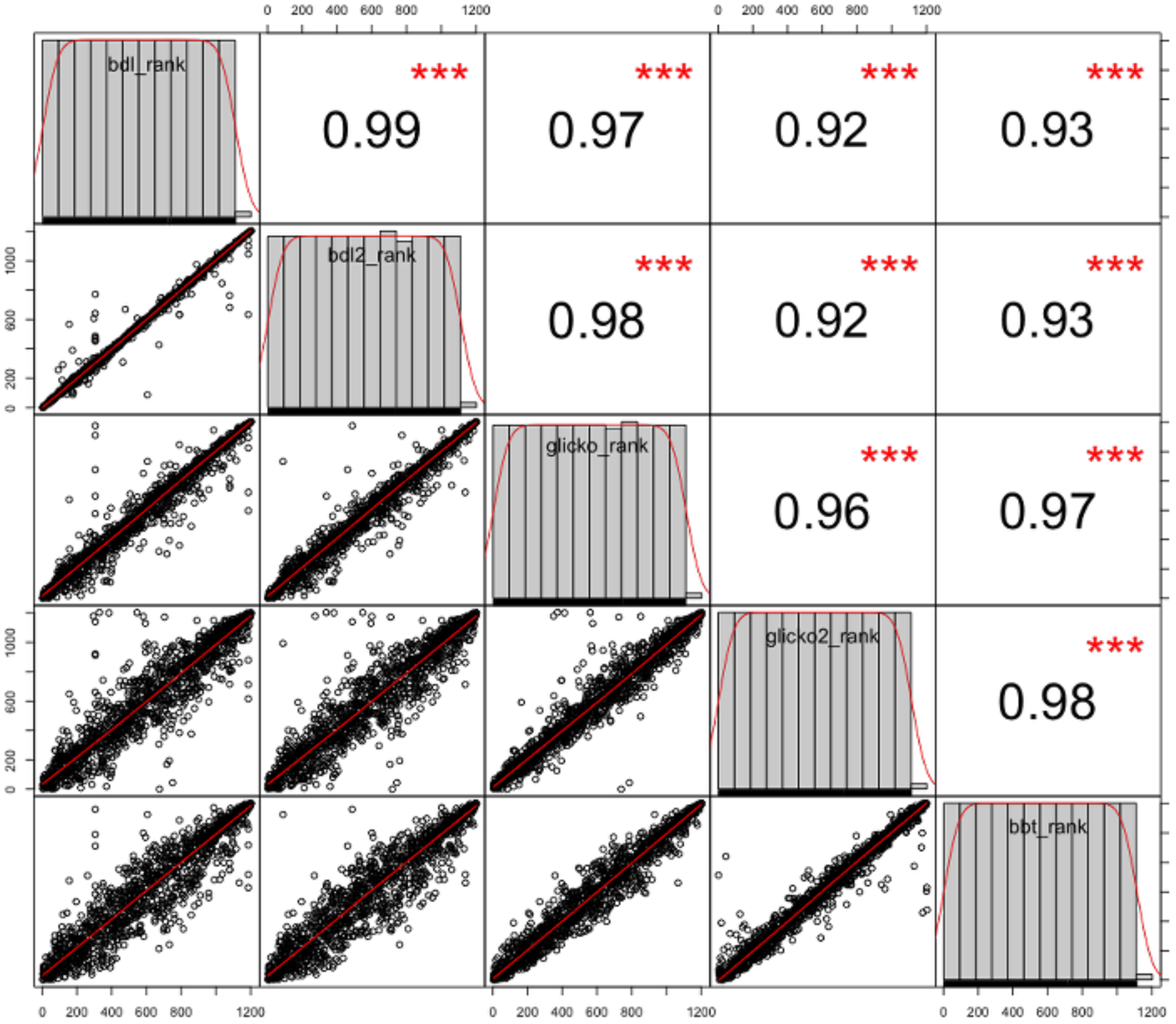
`sport` - package for sport analytics.



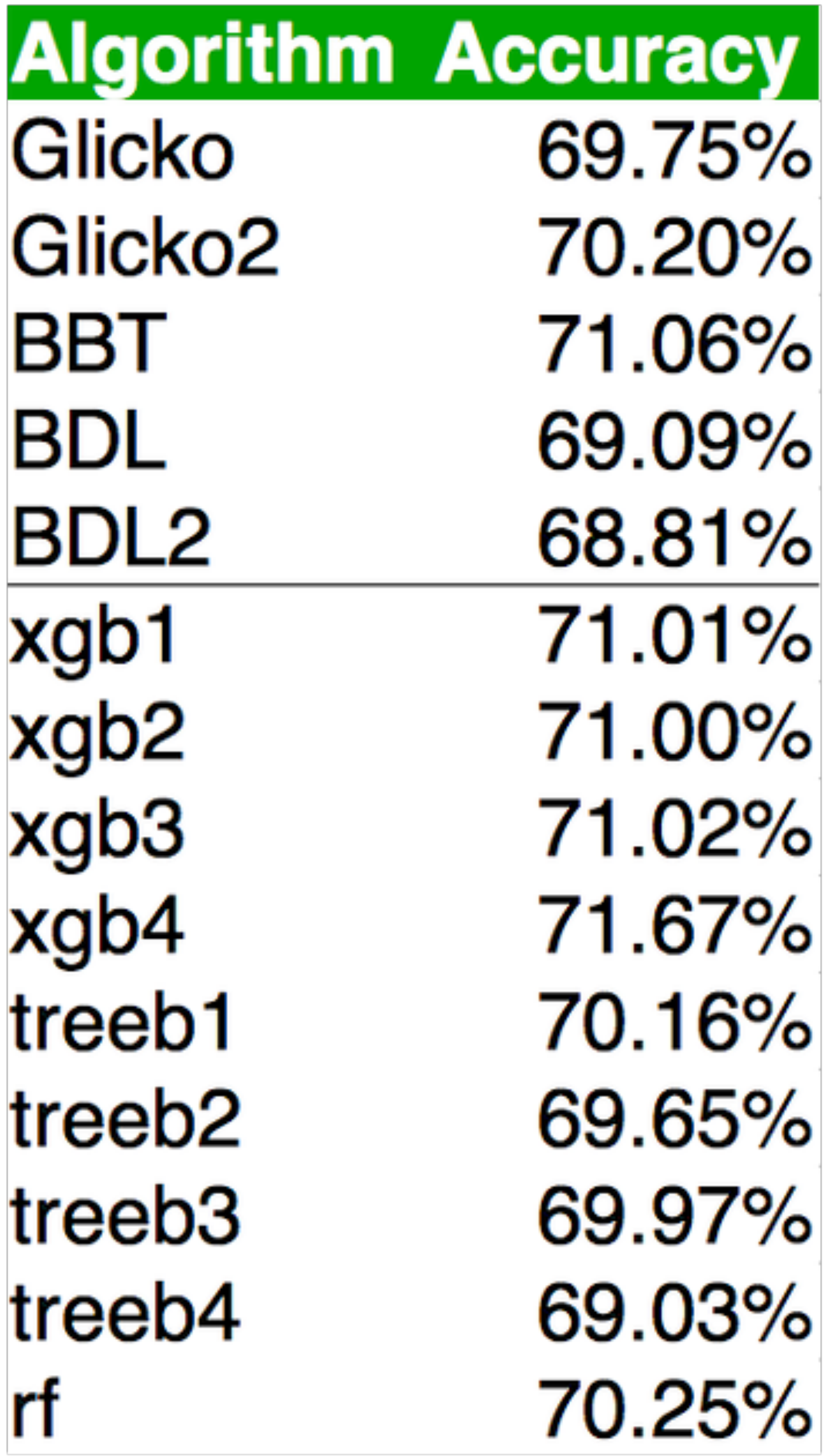
```
# devtools::install_github("gogonzo/sport")
library(sport)

list_glicko  <- glicko_run(  formula = rank|id ~ rider_name, data = gpheats )
list_glicko2 <- glicko2_run( formula = rank|id ~ rider_name, data = gpheats )
list_bbt     <- bbt_run(     formula = rank|id ~ rider_name, data = gpheats )
list_bdl     <- bdl_run(     formula = rank|id ~ rider_name, data = gpheats )
> names(list_glicko)
[1] "r"        "pairs"    "final_r"  "final_rd"
> head(list_glicko$r)
  id           names        r       rd
1  1     Tomasz Gollob 1586.327 203.5029
2  1     Gary Havelock 1241.019 203.5029
3  1       Chris Louis 1758.981 203.5029
4  1 Tony Rickardsson 1413.673 203.5029
5  2     Sam Ermolenko 1758.981 203.5029
6  2    Jan Staechmann 1241.019 203.5029
> tail(list_glicko$pairs)
        id          team1            team2         P Y
61043 5063    Tai Woffinden Fredrik Lindgren 0.6333719 0
61044 5063    Tai Woffinden     Patryk Dudek 0.4877610 0
61045 5063    Tai Woffinden Emil Sajfutdinow 0.5177363 0
61046 5063 Emil Sajfutdinow Fredrik Lindgren 0.6167259 0
61047 5063 Emil Sajfutdinow     Patryk Dudek 0.4700839 0
61048 5063 Emil Sajfutdinow    Tai Woffinden 0.4822637 1
```

Methods perform similarly.



ML methods don't improve model.

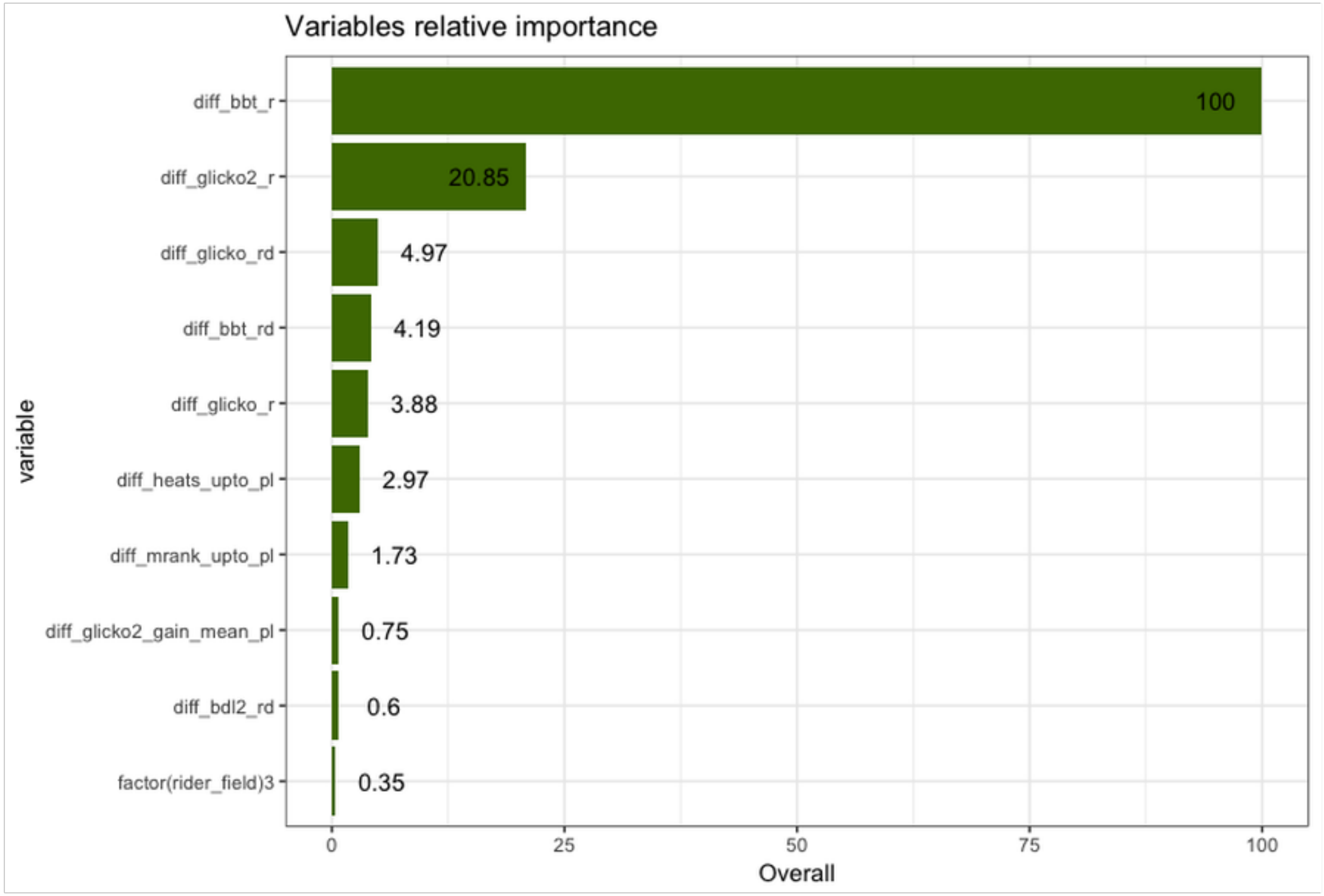


Algorithm	Accuracy
Glicko	69.75%
Glicko2	70.20%
BBT	71.06%
BDL	69.09%
BDL2	68.81%
xgb1	71.01%
xgb2	71.00%
xgb3	71.02%
xgb4	71.67%
treeb1	70.16%
treeb2	69.65%
treeb3	69.97%
treeb4	69.03%
rf	70.25%

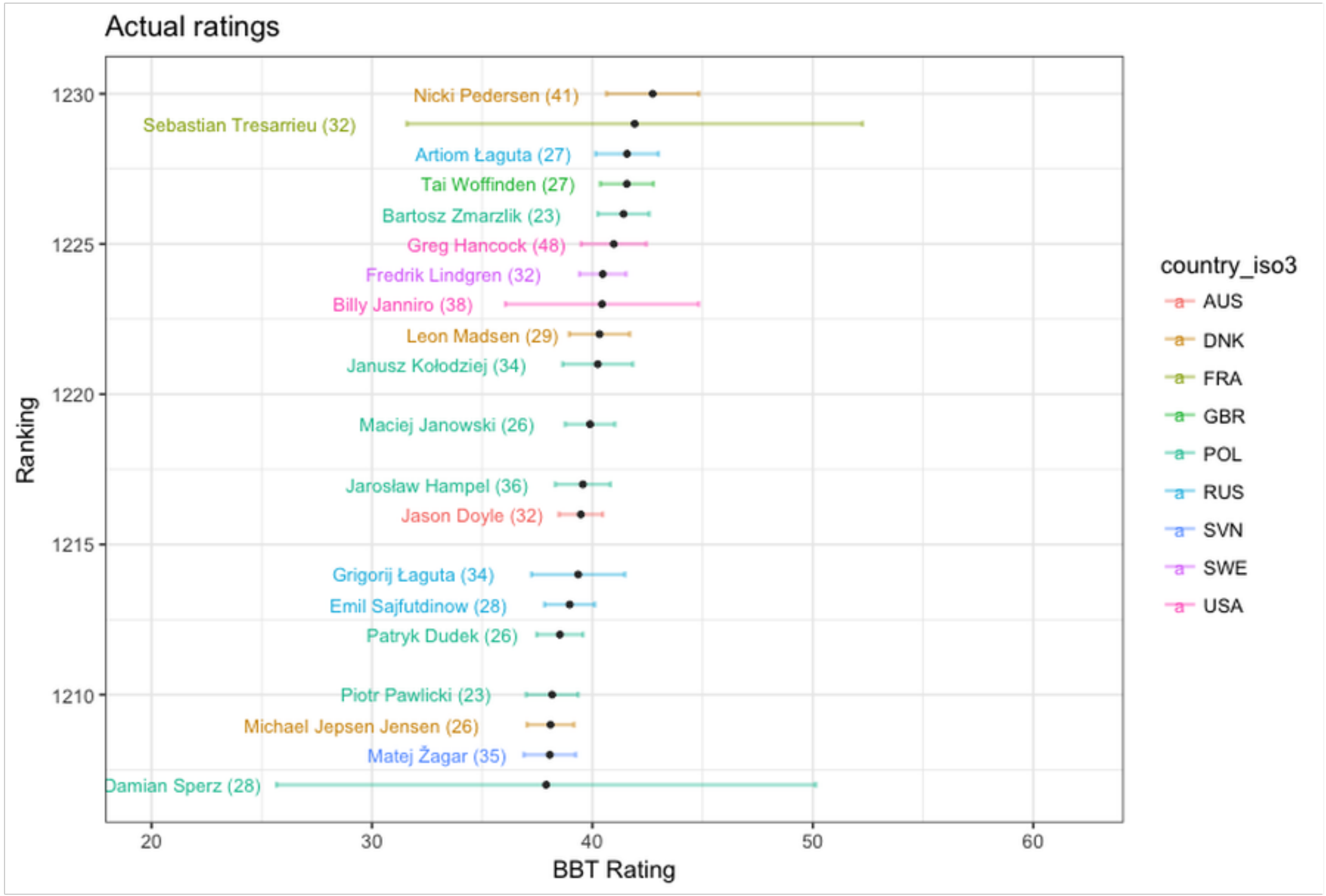
- xgBoost, Random Forests nor Boosted trees didn't improve accuracy.

- Additional variables have no additional predictive abilities.

Non-rating vars contribution < 5%



The Best speedway riders.



What comes next

- Searching for perfect predictive model.

- Improve `sport` package.

- Simulate events results.

- Promote data-science in speedway.

References

- Mark E. Glickman (1999): Parameter estimation in large dynamic paired comparison experiments. Applied Statistics, 48:377-394.
URL http://www.glicko.net/research/glicko.pdf

- Mark E. Glickman (2001): Dynamic paired comparison models with stochastic variances, Journal of Applied Statistics, 28:673-689.
URL http://www.glicko.net/research/dpcmsv.pdf

- Mark E. Glickman (1995): A Comprehensive guide to chess ratings. American Chess Journal, 3, pp. 59--102.
URL http://www.glicko.net/research/acjpaper.pdf

- Ruby C. Weng and Chih-Jen Lin (2011): A Bayesian Approximation Method for Online Ranking. Journal of Machine Learning Research,12:267-300.
URL http://jmlr.csail.mit.edu/papers/volume12/weng11a/weng11a.pdf

- William D. Penny and Stephen J. Roberts (1999): Dynamic Logistic Regression, Departament of Electrical and Electronic Engineering, Imperial College

Thank you

github.com/gogonzo
dawid.kaledkowski@clickmeeting.com
linkedin.com/in/dawidkaledkowski

WHY R
WROCŁAW 2018