

Dr. James Hopker

Using retrospective analysis of race results
to determine success in elite cycling



Aims

- What does the career of a successful cyclist look like?
- Can we determine career prospects from retrospective data analysis?
- Lots of data in public repositories
- Largely unused
- Aim is to give some examples of data utility and point out some difficulties of doing so

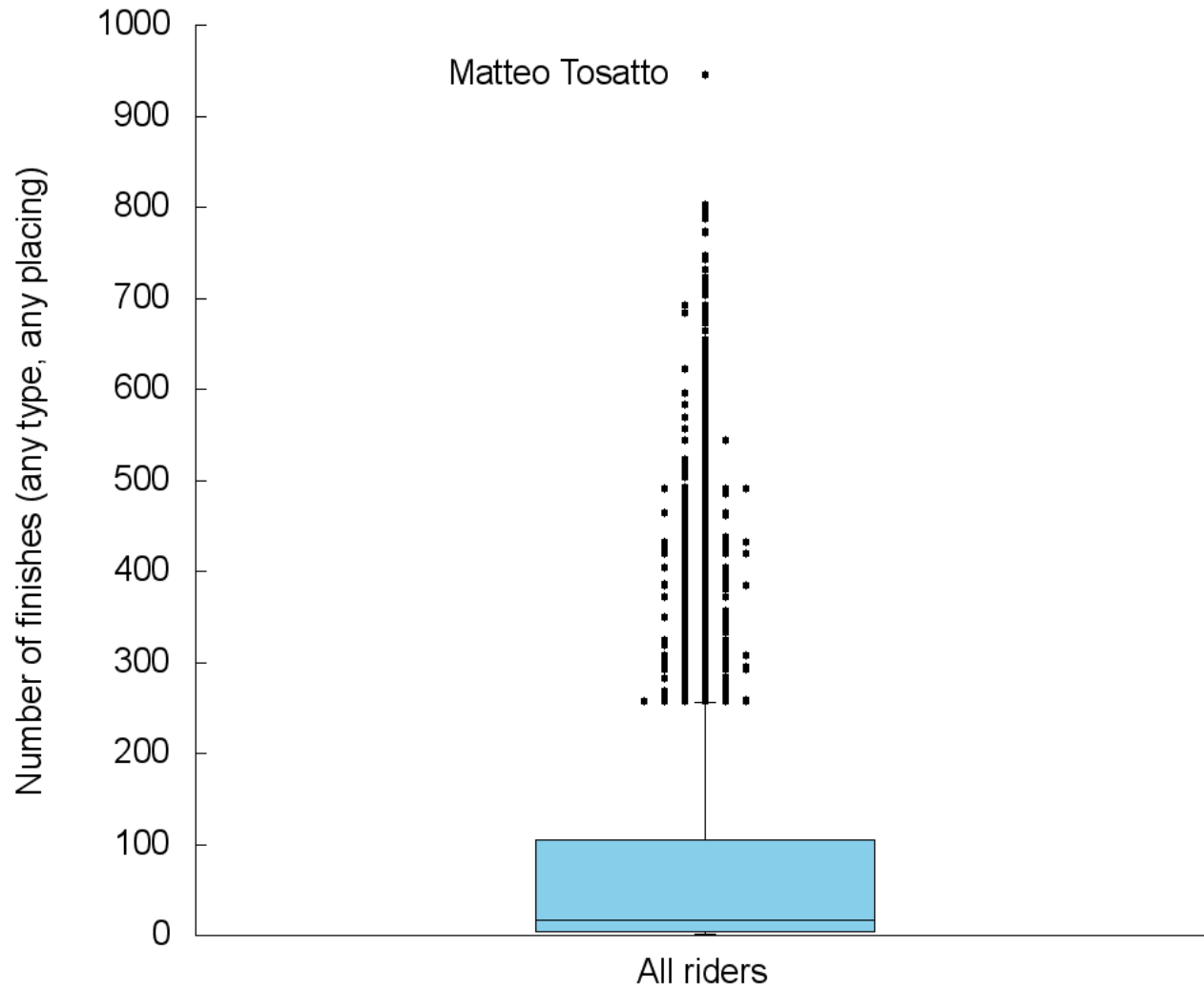
Data details

- Taken from proccyclingstats.com, with permission
- 25 years: 1980 - 2014 inclusive
- Up to 25 races, including multi-stage races
- Results for:
 - One day races
 - Individual stages of multi-day “Tour” races
 - GCs of multi-day stage races
- Total riders: 5928
 - 408,836 results
 - 78 different nationalities

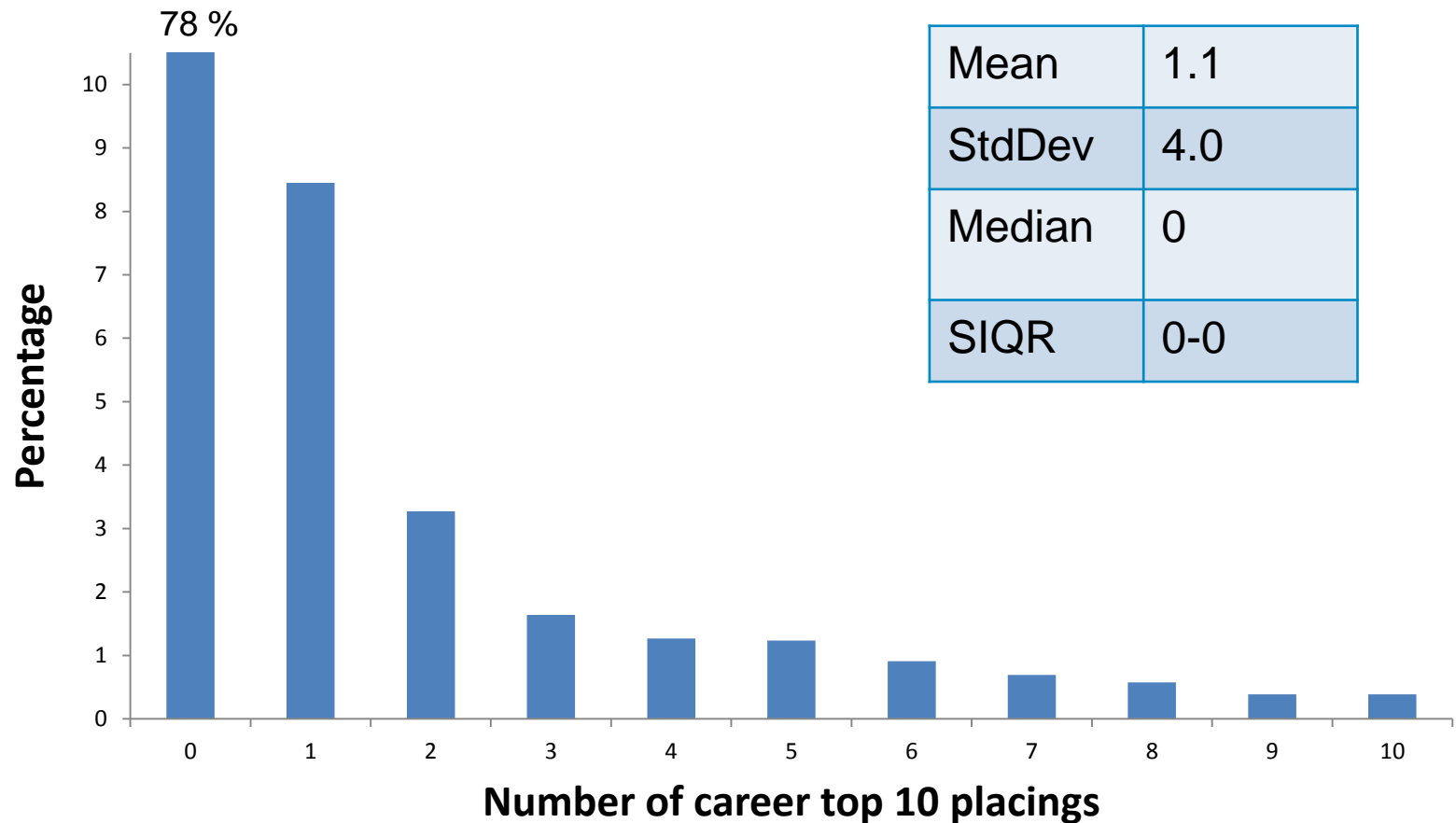
Races included within database

1	Course De La Paix Juniors
2	Paris Roubaix Juniors
3	World Championships U23 Road Race
4	Amstel Gold Race
5	Criterium du Dauphine
6	Eneco Tour
7	Gent Wevelgem
8	Giro d Italia
9	Giro di Lombardia
10	La Fleche Wallonne
11	Liege Bastogne Liege
12	Milano San Remo
13	Paris Nice
14	Paris Roubaix
15	Ronde van Vlaanderen
16	Tirreno Adriatico
17	Tour de France
18	Tour de Pologne
19	Tour de Romandie
20	Tour de Suisse
21	Tour Down Under
22	Volta a Catalunya
23	Vuelta a Espana
24	Vuelta a Pais Vasco
25	World Championships Road Race

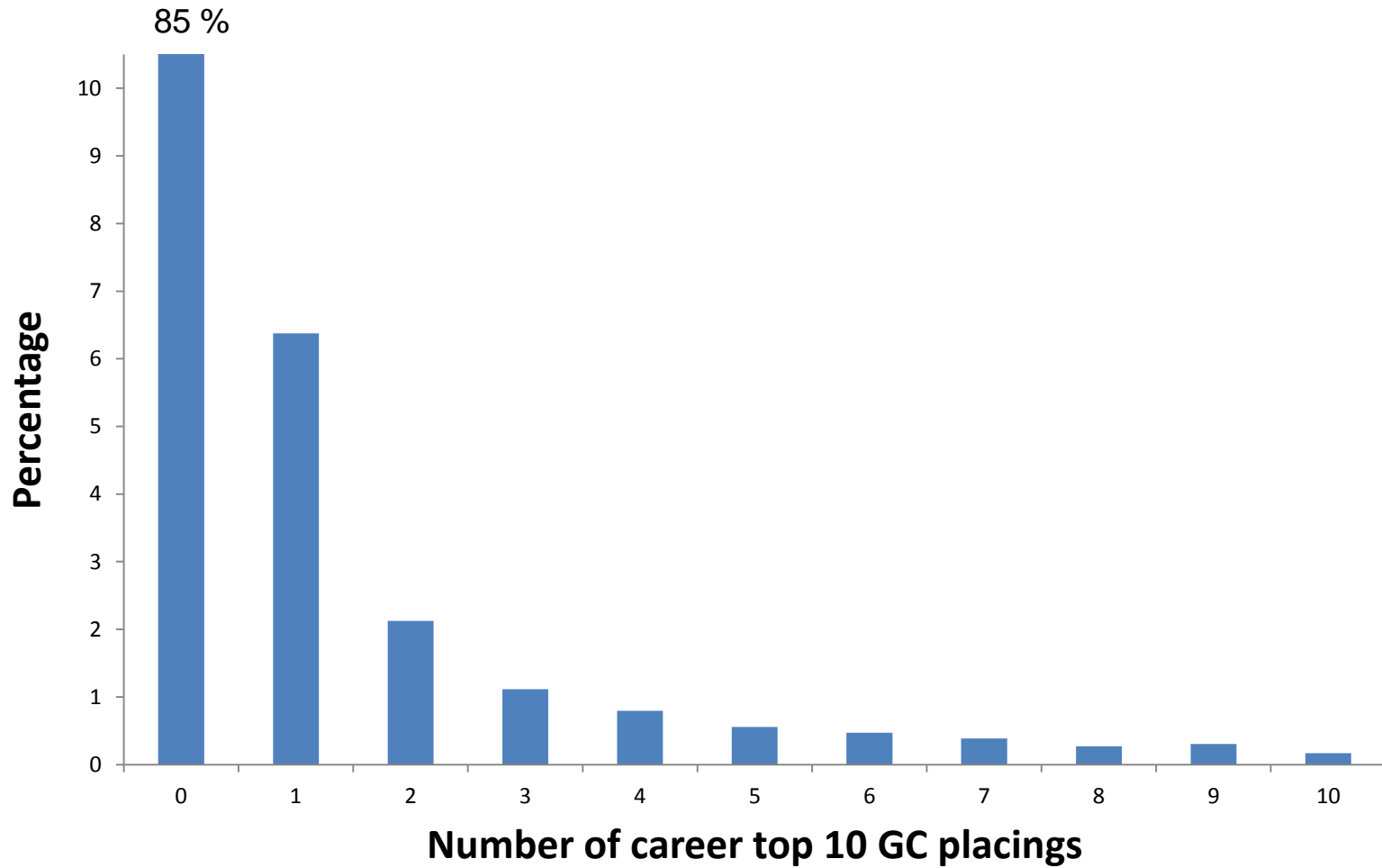
Frequency of Career Results



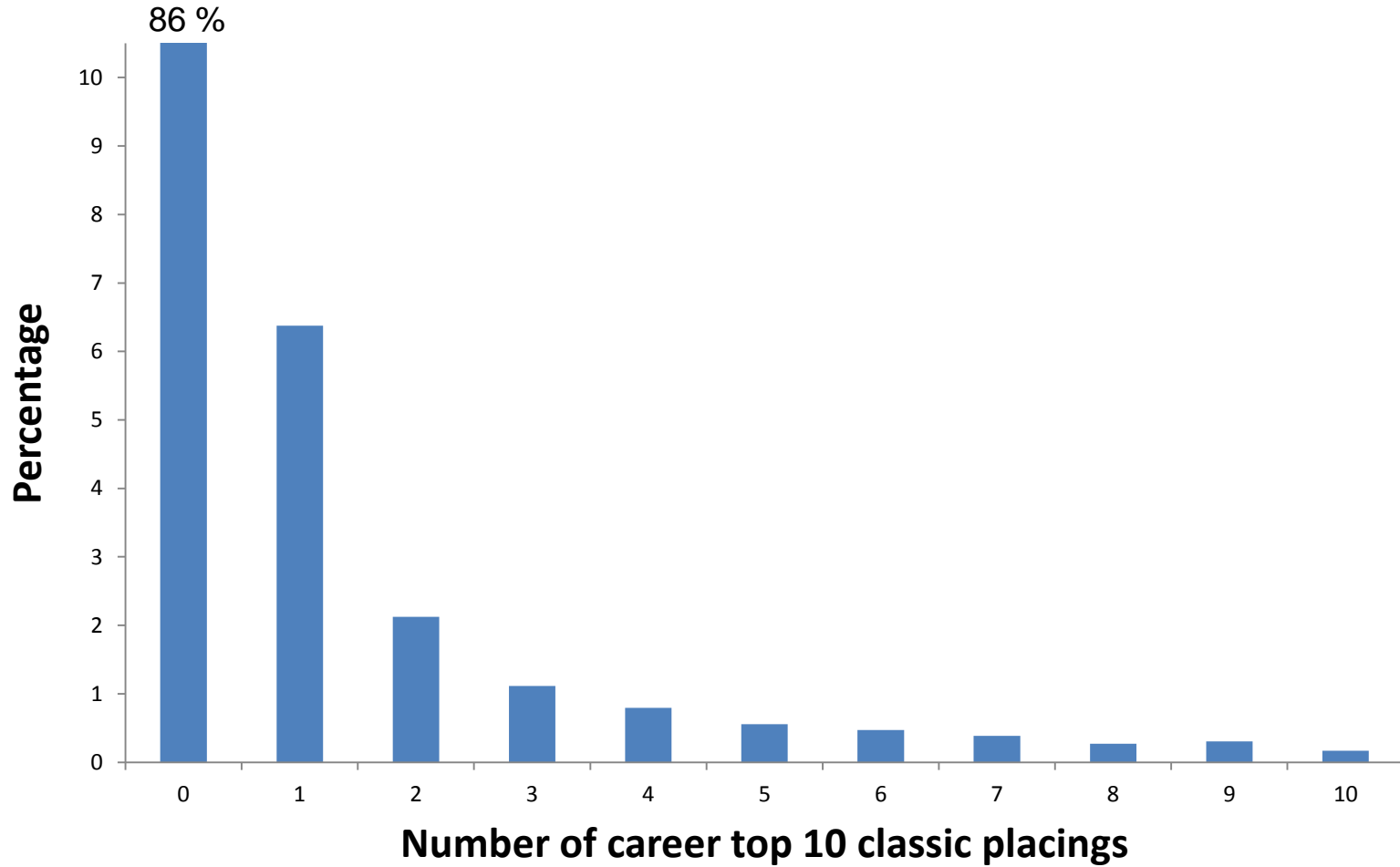
Frequency of top 10 career placings per rider



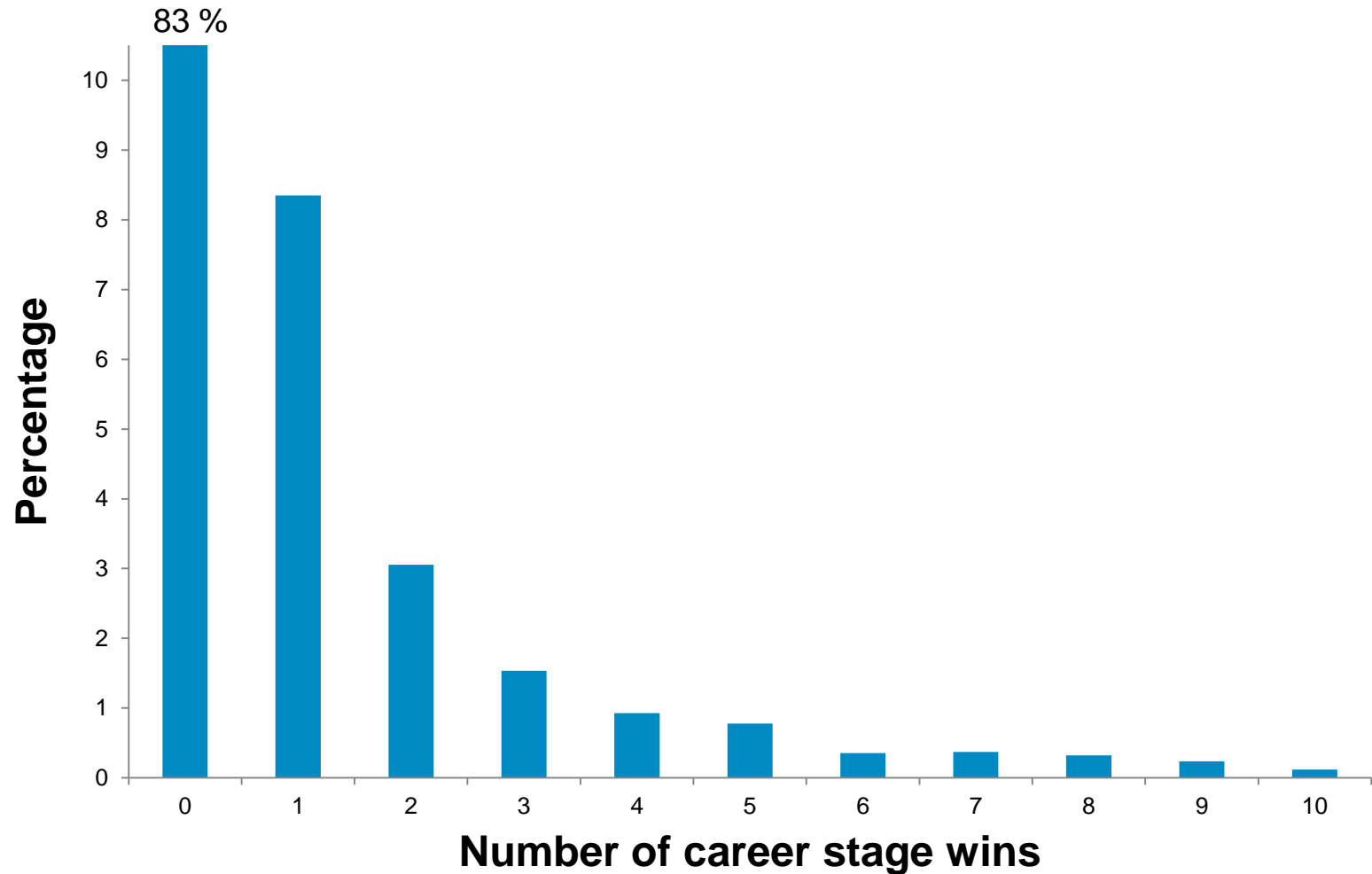
Frequency of top 10 GC placings



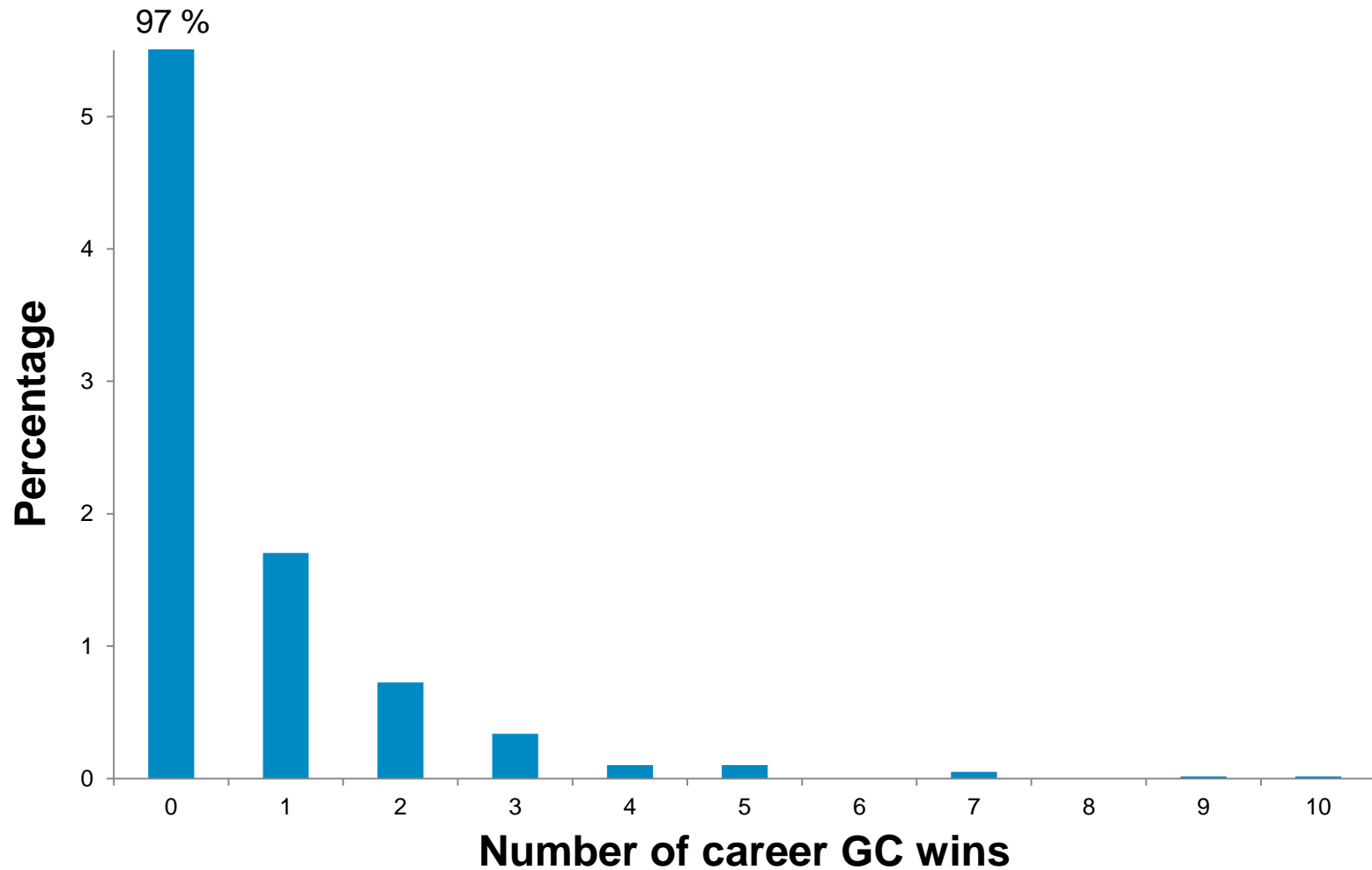
Frequency of classic top 10 placings



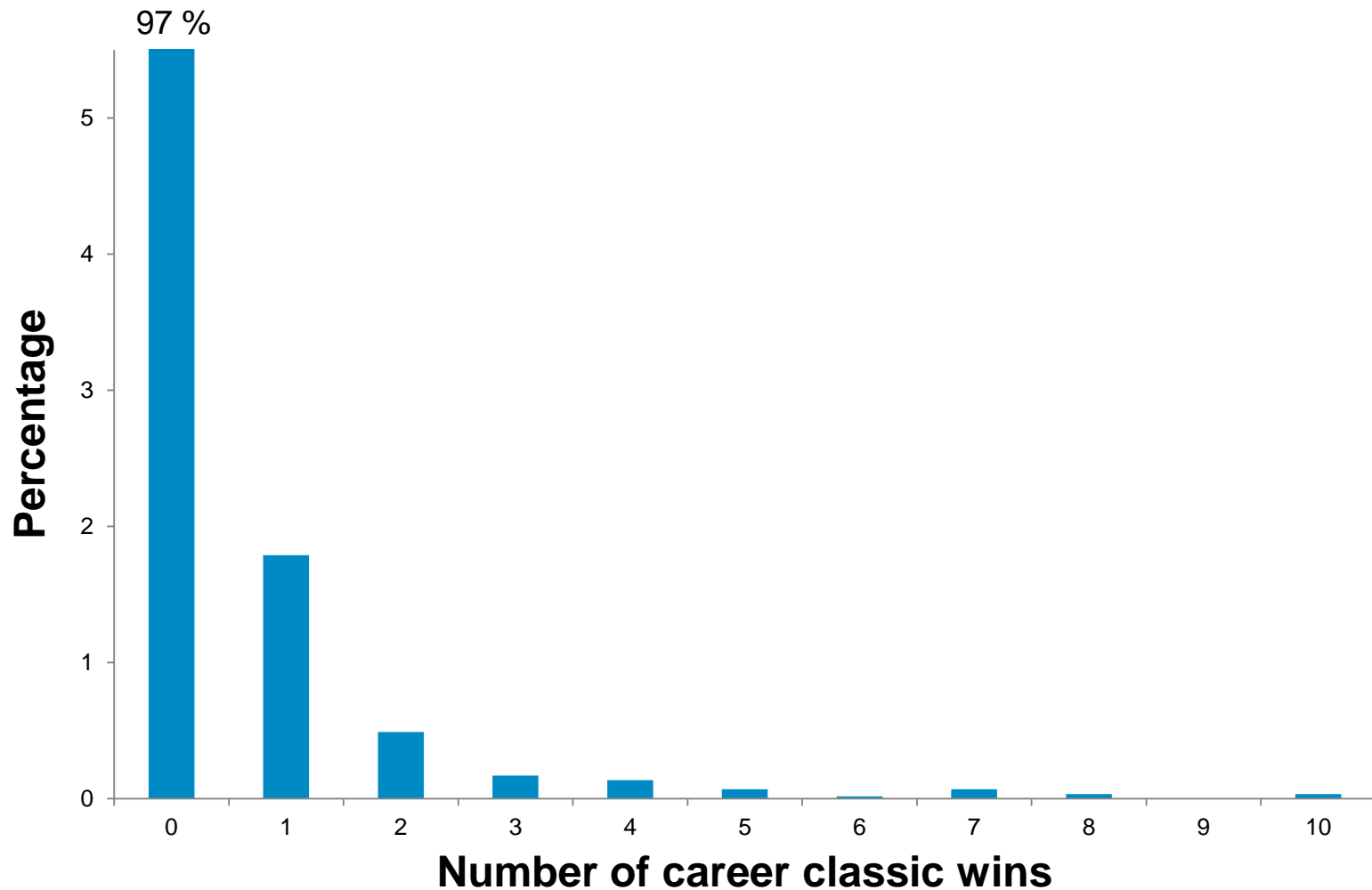
Frequency of Tour Stage Wins per rider



Frequency of GC Wins per rider



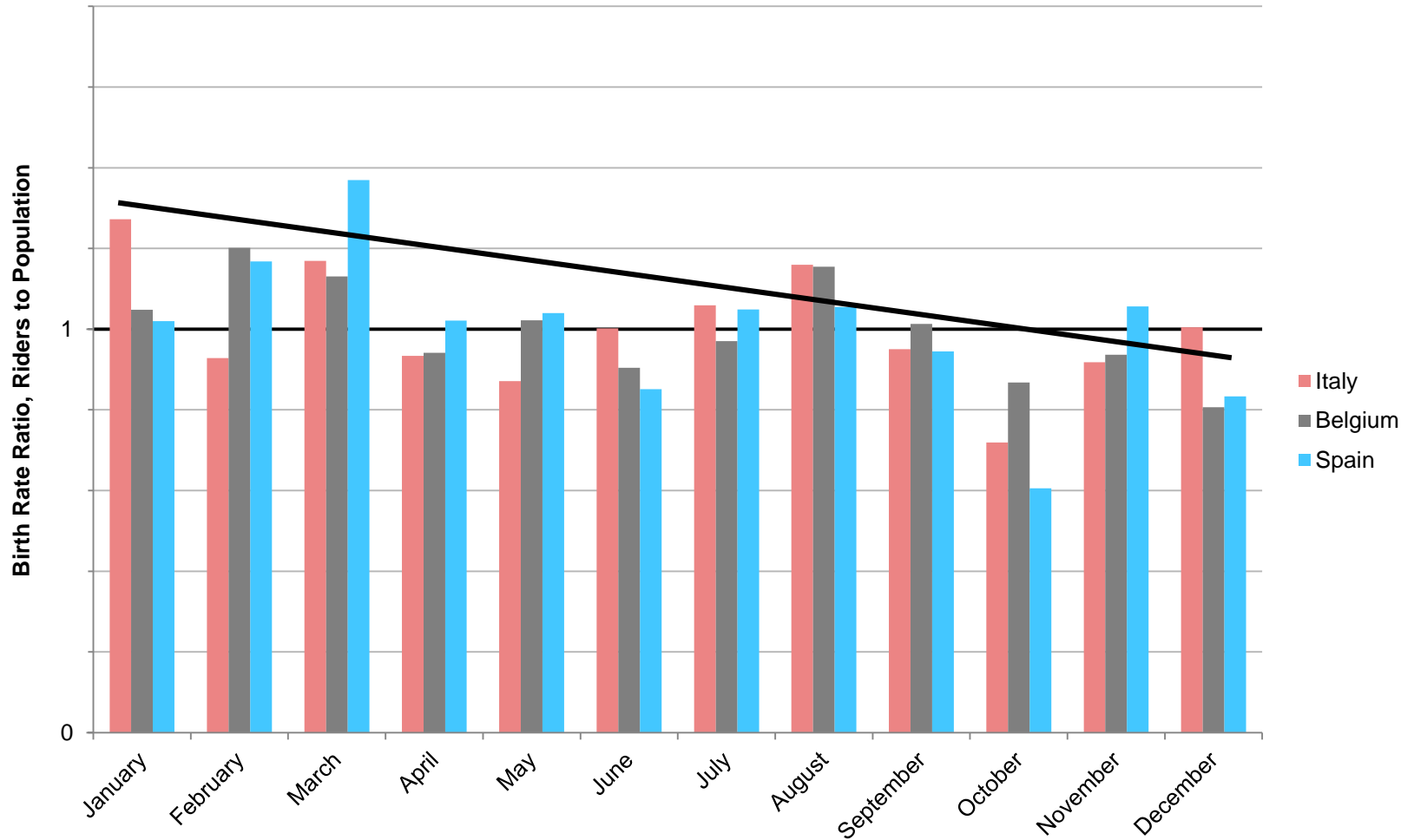
Frequency of one day race Wins per rider



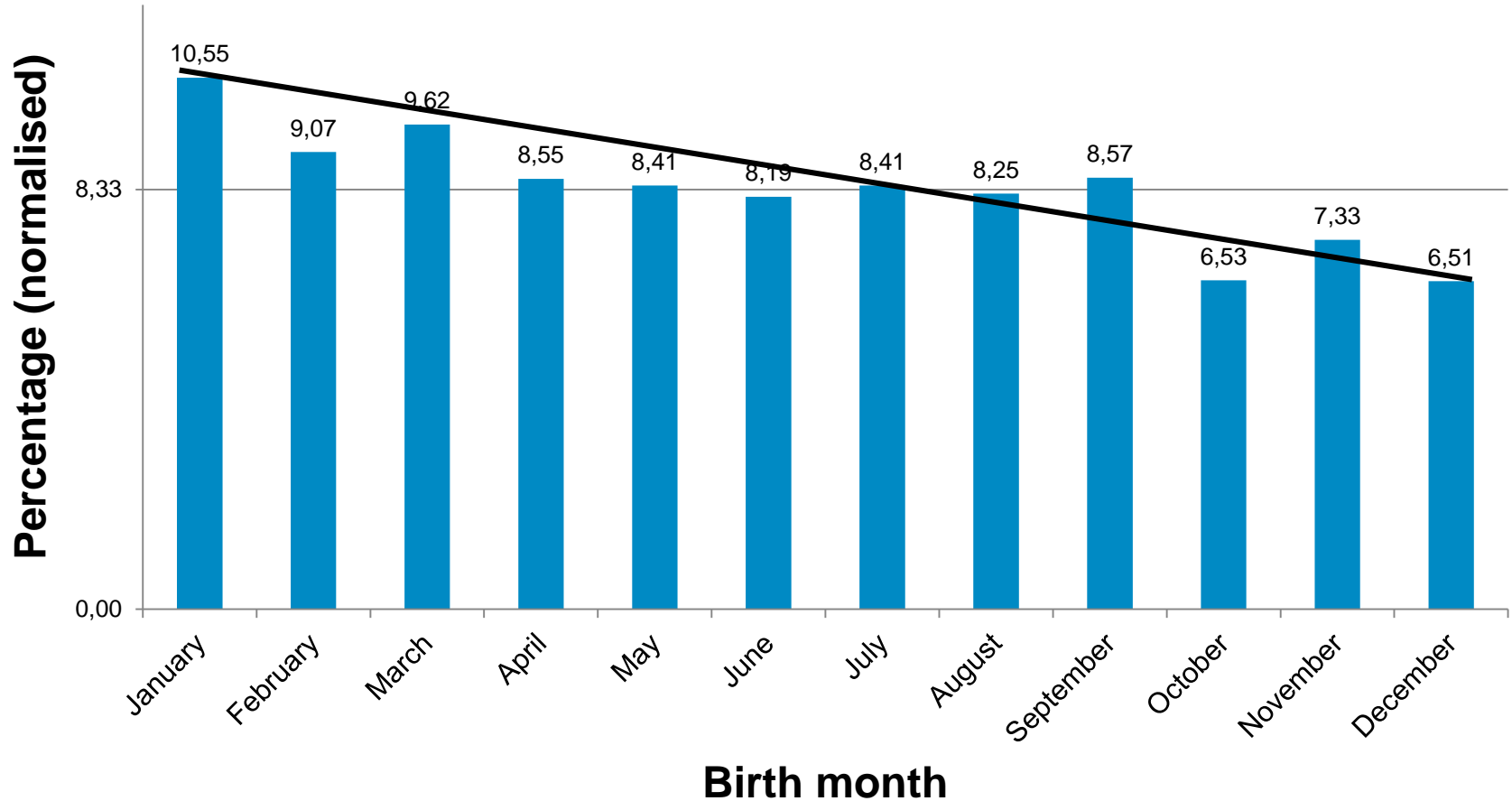
The Relative Age Effect

- Rider selection at end of year
- Possible bias within selection
- Selected initially due to physical, behavioural or psychological maturity advantage, then self-reinforcing
- Issues: normalisation
 - Length of month
 - Birth rate

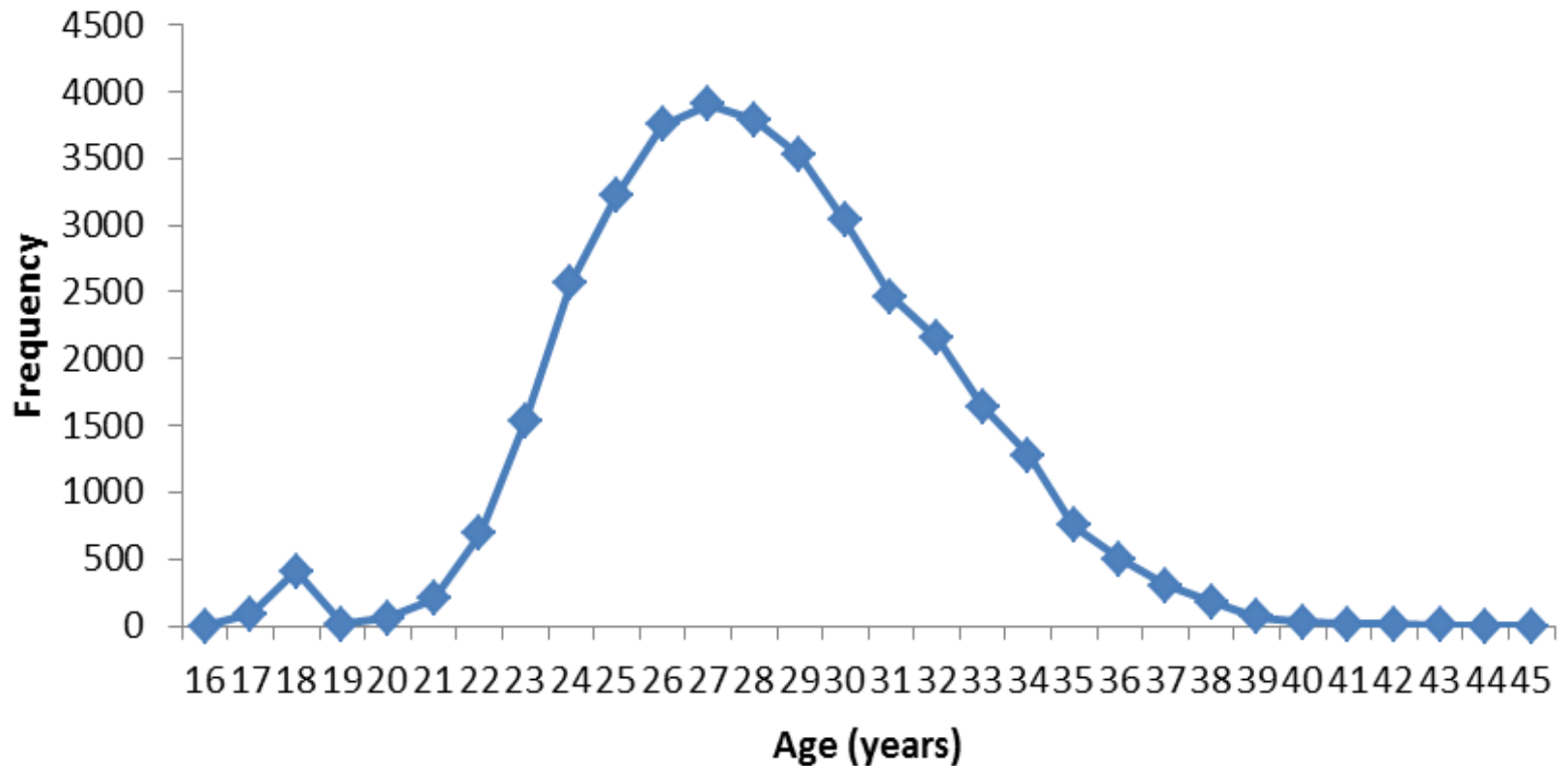
Riders by birth month, compared to country's birth rate



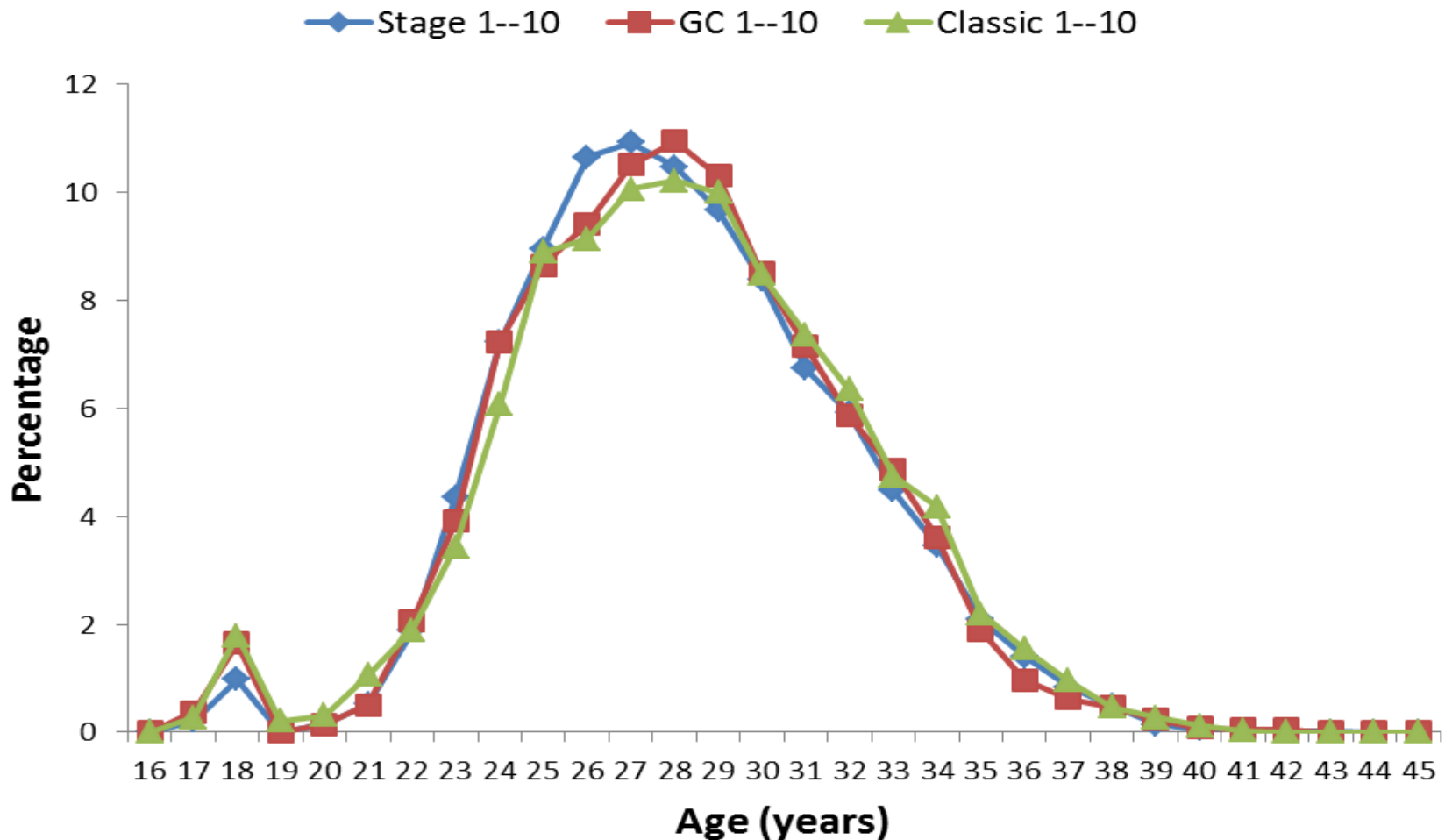
Number of riders placing top 10 by birth month (normalised for month length)



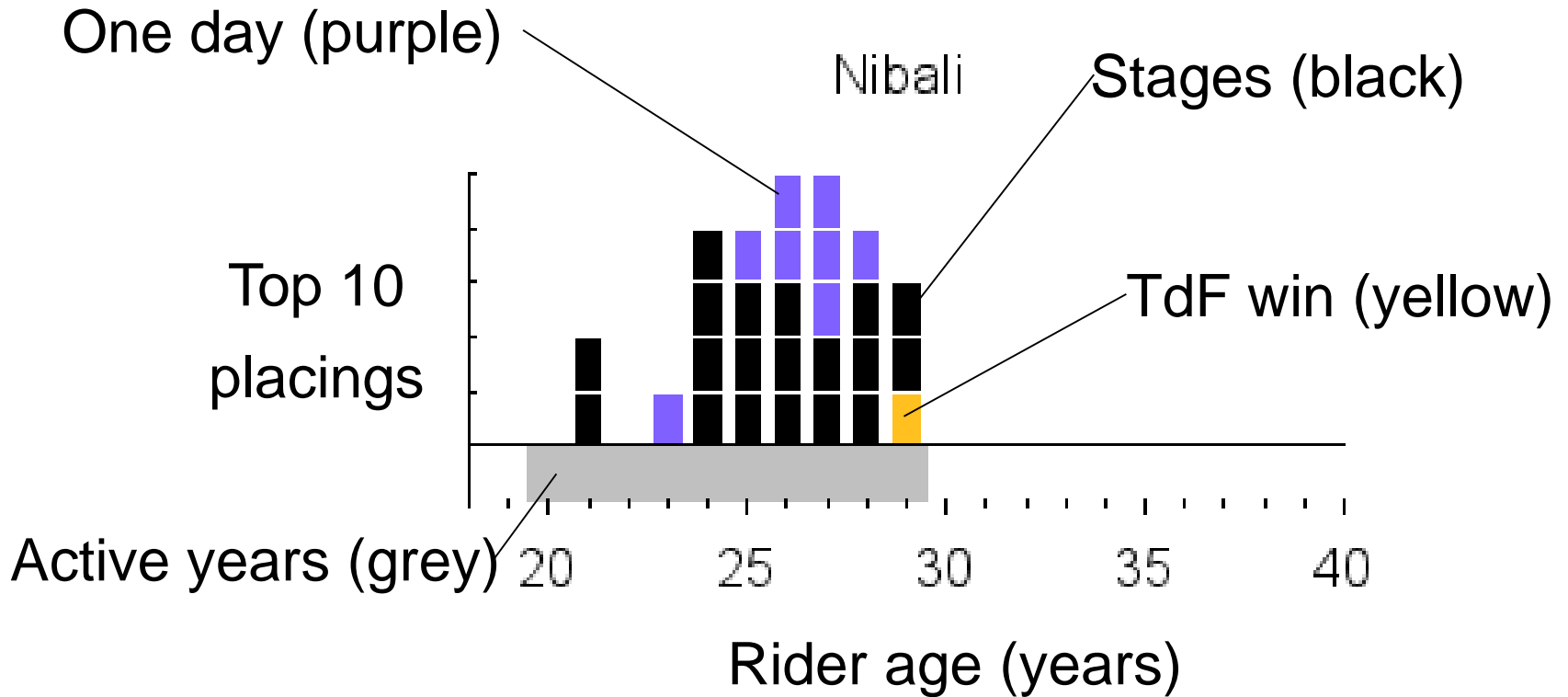
Number of top 10 placings by rider age

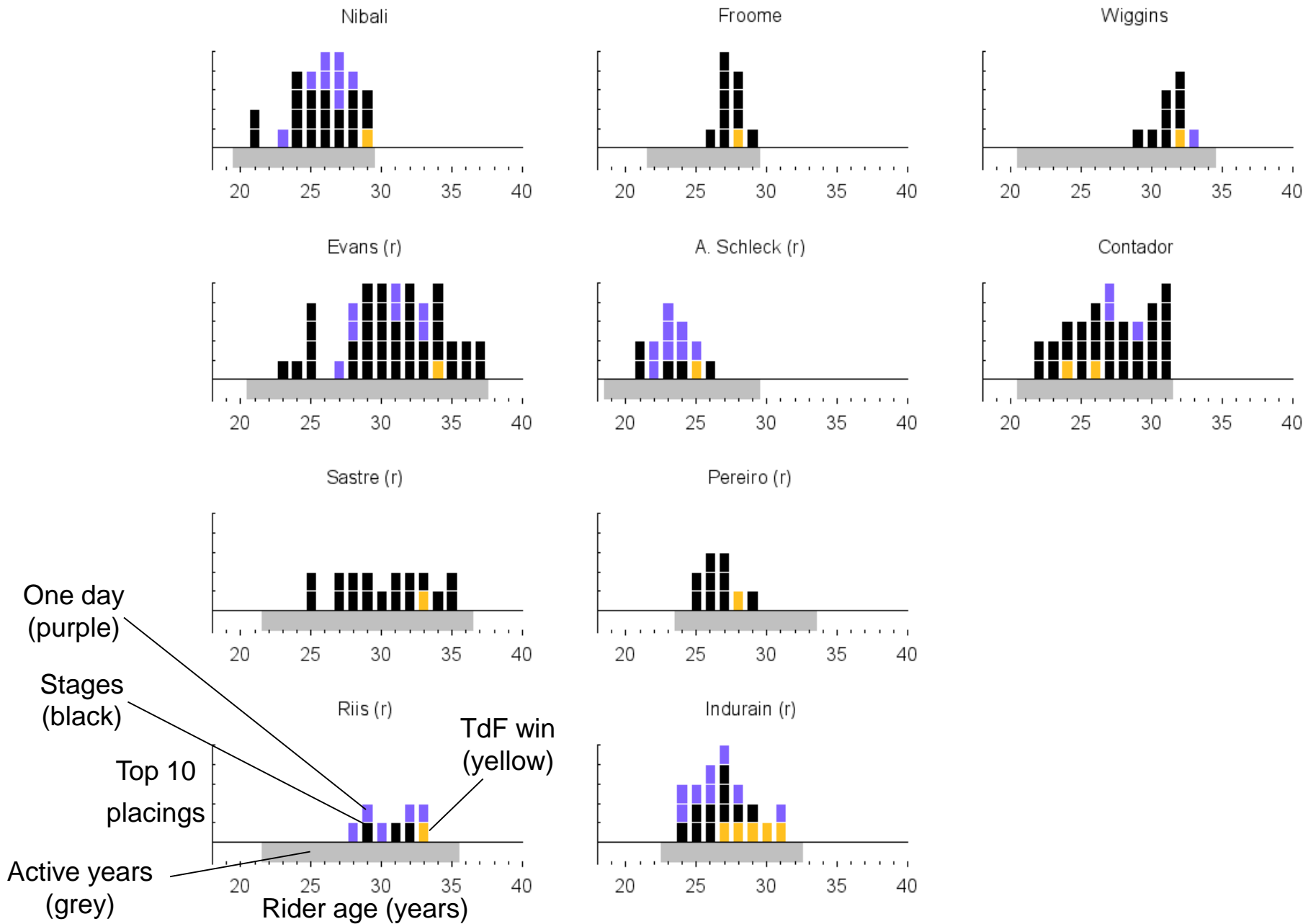


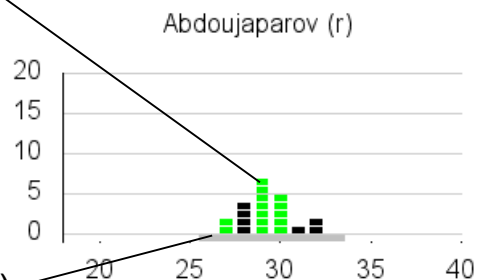
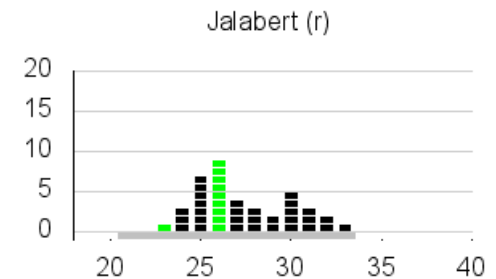
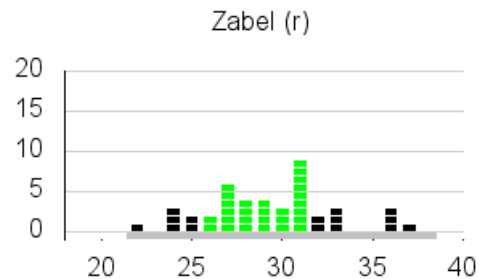
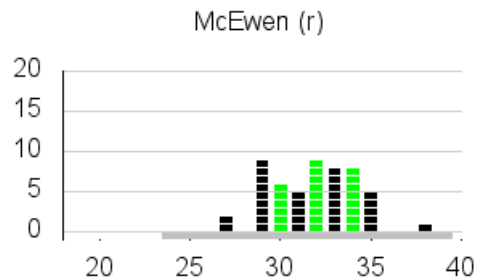
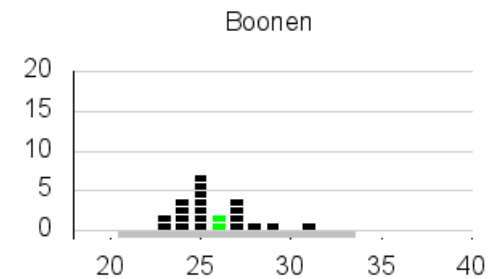
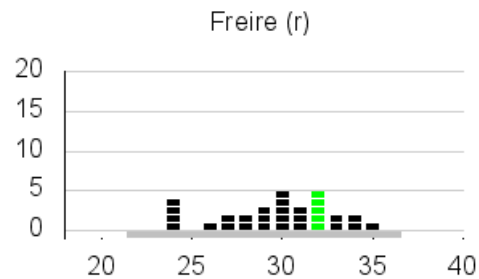
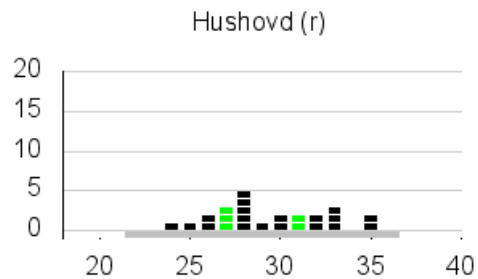
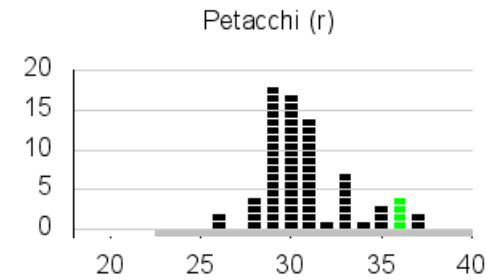
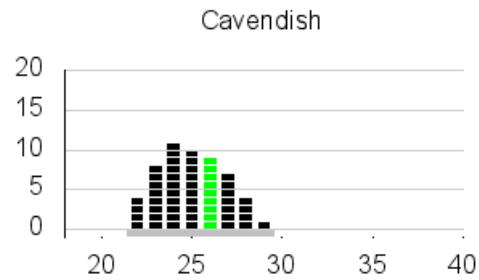
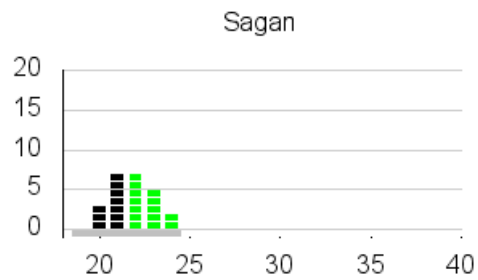
Number of top 10 placings by rider age



What does success look like?







Green: TdF points
win that year

Stage wins

Active years (grey)

Rider age (years)

Conclusions

- 5928 riders, 408,836 results
- Median career of 3 seasons, SIQR: [1, 7]
min. 1, max. 22
- Median of 17 top races completed over career
 - but heavily influenced by a few prolific riders
- Major wins are rare (83 % of riders never win)
 - but database restricted to 25 races, 12 of which are multi-stage races
- Riders biased toward DoB during early months of the year – maturation vs. selection criteria
- Career peak mid to late 20's for most, although some examples of peak occurring ~ 30 years

Acknowledgements

Dr. Kristina Dietz



Prof. Louis Passfield



Prof. Olaf Schumacher



THE UK'S EUROPEAN UNIVERSITY



www.kent.ac.uk

University of
Kent

50

1965-2015
THE UK'S
EUROPEAN
UNIVERSITY