Omega Timing is the official timekeeper for the Olympic Games, including US Olympic Trails. They don’t do very many other events, which is why SwimmeR hasn’t supported Omega-style results. Until now that is. Omega results can now be read into R with versions of SwimmeR >= 0.10.2. We’ll read some Omega results in, and then do a quick set of tests about athlete reaction times.

devtools::install\_github("gpilgrim2670/SwimmeR", build\_vignettes = TRUE)

The 2020 US Trials are being held in 2021, in two parts. Wave I was held June 4th to 7th, and Wave II is currently being held June 13th – 20th.

Let’s get set up and take a look.

library(SwimmeR) library(dplyr) library(stringr) library(ggplot2) library(flextable)

flextable\_style <- function(x) { x %>%

flextable() %>%

bold(part = "header") %>% # bolds header

bg(bg = "#D3D3D3", part = "header") %>% # puts gray background behind the header row

autofit()

}

# US Trials Wave I – Getting Omega Results

The process of reading in Omega results with SwimmeR is exactly the same as reading in Hy- Tek or S.A.M.M.S.. Here’s the entire set of results from Wave I.

file <- "https://github.com/gpilgrim2670/Pilgrim\_Data/raw/master/Omega/Omega\_OT\_Wave1\_

FullResults\_2021.pdf"

Wave\_I <- file %>% read\_results() %>% swim\_parse(splits = TRUE)

Here’s the top three finishers in the Women’s 100 Fly Final. The usual information is present – Place, Name, Team Finals\_Time (Omega results don’t include prelims times…), various Splits columns. Also present is a Reaction\_Time column, that will be the focus of a little demonstration later on.

Wave\_I %>%

filter(Event == "6 JUN 2021 - 7:37 PM Women's 100m Butterfly Final")

%>%

head(3) %>%

select(where( ~ !all(is.na(.)))) %>% # remove splits columns that aren't relevant to this race (Split\_150 etc.)

select(-DQ,

-Exhibition,

"Reaction" = "Reaction\_Time", "Finals" = "Finals\_Time") %>%

flextable\_style()

Place Lane Name Team Reaction Finals Event Split\_50 Split\_100

1 6 LU Sydney PLS 0.64 1:00.38

6 JUN 2021 – 7:37 PM

Women’s 100m Butterfly Final

28.54 31.84

SMITHWICK

2 4

Heidi

JDST 0.68 1:00.56

6 JUN 2021 – 7:37 PM

Women’s 100m Butterfly Final

28.08 32.48

3 5 VANNOTE Ellie UNC 0.69 1:00.60

6 JUN 2021 – 7:37 PM

Women’s 100m Butterfly Final

28.33 32.27

# US Trials Wave II

Wave II of the US trials is where the actual Olympic Team is being selected. It’s still underway as of this writing, so there’s not a single document containing all results available. Individual result documents for each event are being posted however, as the events are completed. Here’s the Women’s 100 Breaststroke final, featuring Lilly King.

file <- "https://github.com/gpilgrim2670/Pilgrim\_Data/raw/master/Omega/Omega\_OT\_Wave2\_

W100Br\_Finals\_2021.pdf"

W100Br <- file %>% read\_results() %>% swim\_parse(splits = TRUE)

W100Br %>%

select(-DQ,

-Exhibition,

"Reaction" = "Reaction\_Time", "Finals" = "Finals\_Time") %>%

flextable\_style()

Place Lane Name Team Reaction Finals Event Split\_50 Split\_100

Place Lane Name Team Reaction Finals Event Split\_50 Split\_100

1. 4 KING Lilly ISC 0.65 1:04.79

PM Women’s 100m Breaststroke Final

30.34 34.45

1. 3 JACOBY Lydia STSC 0.63 1:05.28

PM Women’s 100m Breaststroke Final

30.94 34.34

1. 5 LAZOR Annie MVN 0.66 1:05.60

PM Women’s 100m Breaststroke Final

30.82 34.78

1. 6 GALAT Bethany AGS 0.53 1:05.75

PM Women’s 100m Breaststroke Final

30.69 35.06

DOBLER

5 0

Kaitlyn

TDPS 0.65 1:06.29

PM Women’s 100m Breaststroke Final

30.83 35.46

SUMRALL

6 2

Micah

GAME 0.71 1:06.84

PM Women’s 100m Breaststroke Final

31.83 35.01

7 7 HANNIS Molly TNAQ 0.70 1:07.26

PM Women’s 100m Breaststroke Final

31.29 35.97

ESCOBEDO

8 1

Emily

COND 0.68 1:07.31

PM Women’s 100m Breaststroke Final

31.91 35.40

TUCKER

9 8

Miranda

UN-MI 0.68 1:07.44

PM Women’s 100m Breaststroke Final

31.73 35.71

# Australian Trials

Also underway are the Australian Trials. Similarly to the US Trials they can be read into R using SwimmeR versions >= 0.10.2. For the very curious, these are Hy-Tek results, not Omega. Here’s just the Men’s 100 Fly Final.

file <- "<http://liveresults.swimming.org.au/SAL/2021TRIALS/210612F015.htm>"

M100Bk <- file %>% read\_results() %>% swim\_parse(splits = TRUE)

M100Bk %>%

select(-DQ,

-Exhibition,

-Points,

"Prelims" = "Prelims\_Time", "Finals" = "Finals\_Time") %>%

flextable\_style()

Place Name Age Team Prelims Finals Event Split\_50 Split\_100

1. LARKIN, MITCH 27 STPET 53.04 53.40

Male 100 LC Metre Backstroke

25.86 27.54

1. COOPER, ISAAC 17 RACKL 53.79 53.49

Male 100 LC Metre Backstroke

25.94 27.55

1. HOLLARD, TRISTA 24 STHPT 54.56 54.00

Male 100 LC Metre Backstroke

26.73 27.27

WOODWARD,

4

BRADL

22 MING 54.47 54.13

Male 100 LC Metre Backstroke

26.19 27.94

1. YANG, WILLIAM 22 LNSC 54.75 54.56

Male 100 LC Metre Backstroke

25.98 28.58

1. MAHONEY, TRAVIS 30 MARI 55.03 55.02

Male 100 LC Metre Backstroke

26.78 28.24

1. VAN KOOL, KAI 19 GUSC 54.68 55.13

Male 100 LC Metre Backstroke

26.38 28.75

1. HARTWELL, TY 20 CHAND 55.00 55.23

Male 100 LC Metre Backstroke

26.66 28.57

1. TYSOE, CAMERON 24 GIND 55.05 54.84

Male 100 LC Metre Backstroke

26.46 28.38

1. MILLS, PETER 24 MBAY 55.04 55.30

Male 100 LC Metre Backstroke

26.74 28.56

1. SWINBURN, STUAR 19 UNSW 55.80 55.65

Male 100 LC Metre Backstroke

27.00 28.65

Place Name Age Team Prelims Finals Event Split\_50 Split\_100

1. BAYLISS, JAMES 17 NCOLL 56.06 55.91

Male 100 LC Metre Backstroke

26.68 29.23

1. BOOTH, SHAYE 20 MING 56.33 55.99

Male 100 LC Metre Backstroke

27.21 28.78

1. DAFF, CONOR 18 MBAY 56.25 56.08

Male 100 LC Metre Backstroke

26.93 29.15

1. FOOTE, NATHAN 20 STAND 56.19 56.33

Male 100 LC Metre Backstroke

27.59 28.74

1. CORNWELL, JYE 24 YERPK 56.03 56.43

Male 100 LC Metre Backstroke

27.17 29.26

# US Trials Wave I Reaction Time Demo

Let’s see if there’s a difference between the reaction times of sprinters, mid distance swimmers and distance swimmers in the US Trials Wave I results. We’ll define anyone who swims 50 or 100m distances as a sprinter, anyone who swims the 800 or 1500m distances as a distance swimmer, and everyone else as mid-distance.

For this analysis We’ll need the Lane, Name, Reaction\_Time and Event columns. The other columns won’t be needed, so I’ll remove them.

We can pull distances out the event names. Note however from the 100 Fly results above that the event names contain more information than we’re perhaps used to seeing. Let’s clean that up.

Wave\_I\_Clean <- Wave\_I %>%

select(Lane, Name, Team, Reaction\_Time, Event) %>% # select only columns of interest

mutate(Event = str\_remove(Event, ".\*(?=(Men)|(Women))")) %>% # remove everything in event names before Men or Women

mutate(Reaction\_Time = as.numeric(Reaction\_Time)) # change type of Reaction\_Time column

Now we can classify swimmers by type.

Wave\_I\_Clean <- Wave\_I\_Clean %>%

group\_by(Name) %>% # determining type by athlete mutate(Type = case\_when(

# encode athlete types based on events swam any(str\_detect(Event, "(1500m)|(800m)"), na.rm = TRUE) == TRUE ~

"Distance",

any(str\_detect(Event, "(100m)|(50m)"), na.rm = TRUE) == TRUE ~ "Sprint",

TRUE ~ "Mid"

)) %>%

mutate(Type = factor(Type, levels = c("Sprint", "Mid", "Distance"))) # type as ordered factor for ggplot later

Let’s look at the distribution of reaction times by swimmer type.

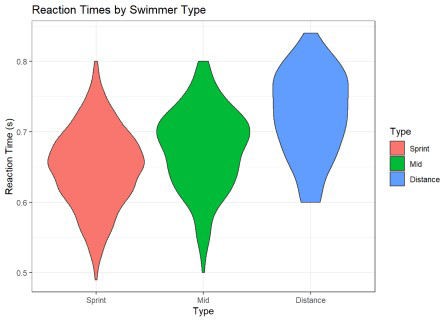
Wave\_I\_Clean %>%

ggplot(aes(x = Type, y = Reaction\_Time, fill = Type)) + geom\_violin() +

theme\_bw() +

labs(y = "Reaction Time (s)",

title = "Reaction Times by Swimmer Type")



There is a noticeable shift towards slower reaction times for distance swimmers compared to sprint and mid-distance, but is it significant? We can use an ANOVA test to determine if the values are significantly different to some standard (called a p value).

reaction\_anova <- aov(Reaction\_Time ~ Type, data = Wave\_I\_Clean) # calculate anova

reaction\_anova\_summary <- summary(reaction\_anova) # save summary anova object

reaction\_anova\_summary # view anova results

## Df Sum Sq Mean Sq F value Pr(>F)

## Type 2 0.479 0.23930 74.65 <2e-16 \*\*\*

## Residuals 1270 4.071 0.00321 ## ---

## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

The p value is very low, at 2.2336931^{-31}. We can conclude that their are significant differences between the groups to at least a significance value (p value) of 0.001. That means the likelihood of these level of difference between the three groups appearing as the result of random variations in populations that are actually identical is less than 0.1%. The ANOVA test doesn’t tell us which group(s) have the significant differences though. For that we can use a

Tukey HSD test.

reaction\_Tukey <- TukeyHSD(reaction\_anova) # calculate Tukey HSD reaction\_Tukey # view results

## Tukey multiple comparisons of means ## 95% family-wise confidence level ##

## Fit: aov(formula = Reaction\_Time ~ Type, data = Wave\_I\_Clean) ##

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ##  ## | $Type | diff | lwr | upr | p | adj |
| ## | Mid-Sprint | 0.02606628 | 0.01762142 | 0.03451114 |  | 0 |
| ## | Distance-Sprint | 0.07474784 | 0.05854508 | 0.09095060 |  | 0 |
| ## | Distance-Mid | 0.04868156 | 0.03158292 | 0.06578019 |  | 0 |

The adjusted p values are all approximately zero. we can see what they actually are by pulling them out of the reaction\_Tukey model object.

reaction\_Tukey$Type[,"p adj"] # view actual adjusted p values ## Mid-Sprint Distance-Sprint Distance-Mid

## 1.634137e-12 0.000000e+00 1.058689e-10

All very low, so all the groups have differences significant at the p = 0.001 level. Sprinters really do have faster reaction times than mid-distance, who are in turn faster than distance swimmers.

# Reaction Times By Lane

Just for giggles let’s also look by lane. When I was swimming there was always this rumor going around that swimmers in the outside lane nearest the starting device would have an advantage, because the light/sound from the device would reach them before it reached athletes further from the device. It never made much sense, since faster swimmers were deliberately seeded into inner lanes and they usually won. Nowadays each block is equipped with a LED light bar and a sounding device so everything should be equal (if it ever wasn’t).

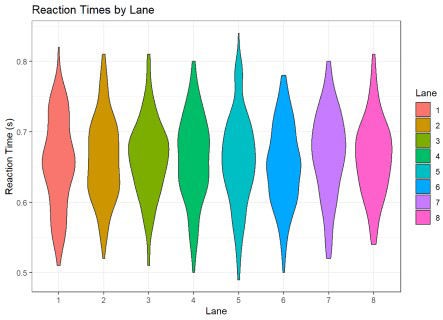
Wave\_I\_Clean %>% filter(Lane != "0") %>%

ggplot(aes(x = Lane, y = Reaction\_Time, fill = Lane)) + geom\_violin() +

theme\_bw() +

labs(y = "Reaction Time (s)",

title = "Reaction Times by Lane")



That looks about even to me. Let’s see what the testing has to say.

reaction\_anova <- aov(Reaction\_Time ~ Lane, data = Wave\_I\_Clean) # calculate anova

reaction\_anova\_summary <- summary(reaction\_anova) # save summary anova object

reaction\_anova\_summary # view anova results

## Df Sum Sq Mean Sq F value Pr(>F)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ## | Lane | 8 | 0.025 | 0.003144 | 0.878 | 0.534 |
| ## | Residuals | 1264 | 4.525 | 0.003580 |  |  |

Here the p value is 0.5341483, which is larger than any p value we’d care to use. There is no significant difference in reaction time by lane.