The first round of the 2020 High School Swimming State-Off Tournament is in the books and saw California (1), Texas (2), Florida, and Pennsylvania (5) advance.

Before beginning the next round there are a few administrative details I’d like to cover.

1. First and foremost: SwimmeR version 0.4.1 is now available on CRAN! The State-Off has been the first major outing for my SwimmeR package. We’ve used it extensively to read in and parse swimming results from a variety of sources, including “normal” html web pages, Hy-Tek real time results pages, and .pdf files. It’s performed admirably, but some bugs have revealed themselves behind the scenes. Version 0.4.1 contains bug fixes plus a host of new features:

A version of results\_score, the function we developed during the State-Off. It handles timed finals style meets (like the State-Off) but also scores prelims-finals style meets, a more common and also more complex format.

library(stringr) library(dplyr) library(purrr) library(SwimmeR) library(flextable)

base <- "<http://sidearmstats.com/auburn/swim/200218F0>" event\_numbers <-

1:42 # sequence of numbers, total of 42 events across men and women event\_numbers <-

str\_pad(event\_numbers,

width = 2, side = "left",

pad = "0") # add leading zeros to single digit numbers SEC\_Links <-

paste0(base, event\_numbers, ".htm") # paste together base urls and sequence of numbers (with leading zeroes as needed)

SEC\_Results <-

map(SEC\_Links, read\_results, node = "pre") %>% # map SwimmeR::read\_results over the list of links

map(

swim\_parse, typo = c(

"A&M",

"FLOR",

"Celaya-Hernande",

# names which were cut off, and missing the last, first structure "Hernandez-Tome",

"Garcia Varela,", "Von Biberstein,"

),

replacement = c( "AM",

"Florida", "Celaya, Hernande",

# replacement names that artificially impose last, first structure. Names can be fixed after parsing

"Hernandez, Tome", "Garcia, Varela", "Von, Biberstein"

)

) %>%

bind\_rows()

# some diving finals results don't list places 9-24, which do score. we can get those divers from the prelim results

SEC\_Diving\_Prelims\_Links <- c(

"<http://sidearmstats.com/auburn/swim/200218P015.htm>", # M 1m prelims "<http://sidearmstats.com/auburn/swim/200218P001.htm>", # W 1m prelims "<http://sidearmstats.com/auburn/swim/200218P022.htm>", # W 3m prelims "<http://sidearmstats.com/auburn/swim/200218P029.htm>", # M platform prelims "<http://sidearmstats.com/auburn/swim/200218P040.htm>"

) # W platform prelims

SEC\_Diving\_Prelims <-

map(SEC\_Diving\_Prelims\_Links, read\_results, node = "pre") %>% # map SwimmeR::read\_results over the list of links

map(

swim\_parse,

typo = c("A&M", "FLOR", "Celaya-Hernande", "Garcia Varela,"), replacement = c("AM", "Florida", "Celaya, Hernande", "Garcia, Varela")

) %>%

bind\_rows()

SEC\_Diving\_Prelims <- SEC\_Diving\_Prelims %>%

anti\_join(SEC\_Results, by = c("Name", "School", "Event")) # make sure divers aren't counted twice for a given event

SEC\_Results <- bind\_rows(SEC\_Results, SEC\_Diving\_Prelims) SEC\_Results <-

SEC\_Results %>% # actual use of new results\_score function results\_score(

events = unique(SEC\_Results$Event), meet\_type = "prelims\_finals", lanes = 8,

scoring\_heats = 3, point\_values = c(

32,

28,

27,

26,

25,

24,

23,

22,

20,

17,

16,

15,

14,

13,

12,

11,

9,

7,

6,

5,

4,

3,

2,

1

)

)

SEC\_Results\_Gender <- SEC\_Results %>%

mutate(Gender = case\_when(str\_detect(Event, "Men") ~ "M",

str\_detect(Event, "Women") ~ "F")) %>% group\_by(School, Gender) %>%

summarise(Score = sum(Points, na.rm = TRUE)) %>% arrange(desc(Score)) %>%

arrange(Gender) %>% ungroup() %>% group\_split(Gender)

The scored results match the official results for women:

SEC\_Results\_Gender[[1]] %>% flextable() %>% bold(part = "header") %>%

bg(bg = "#D3D3D3", part = "header") %>% autofit()

|  |  |  |
| --- | --- | --- |
| **School** | **Gender** | **Score** |
| Tennessee | F | 1108.0 |
| Florida | F | 1079.5 |
| Kentucky | F | 987.5 |
| Georgia | F | 986.0 |
| Auburn | F | 866.0 |
| Texas AM | F | 851.0 |
| Alabama | F | 748.0 |
| Missouri | F | 500.0 |
| South Carolina | F | 427.0 |
| Arkansas | F | 422.0 |
| LSU | F | 417.0 |
| Vanderbilt | F | 150.0 |

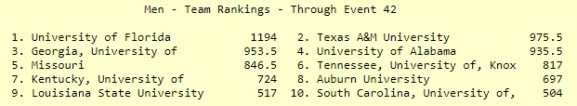


Scores also match for men:

SEC\_Results\_Gender[[2]] %>% flextable() %>% bold(part = "header") %>%

bg(bg = "#D3D3D3", part = "header") %>% autofit()

|  |  |  |
| --- | --- | --- |
| **School** | **Gender** | **Score** |
| Florida | M | 1194.0 |
| Texas AM | M | 975.5 |
| Georgia | M | 953.5 |
| Alabama | M | 935.5 |
| Missouri | M | 846.5 |
| Tennessee | M | 817.0 |
| Kentucky | M | 724.0 |
| Auburn | M | 697.0 |
| LSU | M | 517.0 |
| South Carolina | M | 504.0 |



The ability to read in .hy3 files. Hy-Tek .hy3 files are another form of results, intended to be read into Team Manager. As of version 0.4.1 SwimmeR can now also read them. This feature is *not* complete and will evolve in future releases. Here though we can use it to read in results from the USA Swimming 2019 December Sectional Meet for CA and NV.

temp <- tempfile() temp2 <- tempfile() url <-

"<http://www.pacswim.org/userfiles/meets/documents/1691/meet-results-speedo-sectionals-2019-ca-nv-> december-2019-13dec2019-003.zip"

download.file(url, temp) unzip(zipfile = temp, exdir = temp2) raw\_results <-

read\_results( file.path(

temp2,

"Meet Results-Speedo Sectionals 2019 CA-NV December 2019-13Dec2019-003.hy3"

)

)

unlink(c(temp, temp2))

results <- swim\_parse(raw\_results) %>% mutate(Event = str\_replace(Event, "NA", "Yard"))

results %>%

filter(Event == "100 Yard Butterfly",

Gender == "M") %>%

select(Name, Team = School, Prelims\_Time, Finals\_Time) %>% arrange(Finals\_Time) %>%

head(5) %>% flextable() %>%

bold(part = "header") %>%

bg(bg = "#D3D3D3", part = "header") %>% autofit()

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Team** | **Prelims\_Time** | **Finals\_Time** |
| Fischer, Brandon | C1LAC | 49.20 | 48.07 |
| Antoniuk, Konrad | Paseo Aquatics Swim Team | 50.48 | 50.03 |
| Toland, Brandon | Golden West Swim Club | 50.30 | 50.06 |
| Kim, William | Monterey Park Manta Rays | 50.93 | 50.16 |
| Bowman, Andrew | San Clemente Aquatics | 50.95 | 50.30 |

Recording of DQ and Exhibition swims in the output of swim\_parse, as the columns DQ and Exhibition respectively. This ended up being important for results\_score, since Exhibition and DQ swimmers can’t score.

Ithaca\_Union <- swim\_parse(

read\_results( "https://athletics.ithaca.edu/services/download\_file.ashx?file\_location=https://s3.

amazonaws.com/sidearm.sites/bombers.ithaca.edu/documents/2020/2/1/ithaca\_vs\_union\_2020.pdf"

)

)

Ithaca\_Union %>%

filter(Event == "Men 400 Yard Freestyle Relay") %>% select(Place, School, Finals\_Time, Exhibition, DQ) %>% flextable() %>%

bold(part = "header") %>%

|  |  |  |  |
| --- | --- | --- | --- |
| bg(bg = "#D3D3D3", part autofit() | = "header") | %>% |  |
| **Place School** | **Finals\_Time** | **Exhibition** | **DQ** |
| 1 Ithaca College-NI | 3:21.86 | 0 | 0 |
| 2 Ithaca College-NI | 3:26.28 | 0 | 0 |
| 3 Ithaca College-NI | 3:34.10 | 1 | 0 |
| NA Union College (New York)-MR |  | 0 | 1 |

We can see that in the Mens 400 Yard Freestyle Relay the third place relay was exhibition (Exhibition == 1) and that another relay was disqualified (DQ == 1).

