You can get your very own copy of this cutting edge sports-data-science package, for free, today!

install.packages("JumpeR") library(JumpeR) library(flextable) library(dplyr) library(ggplot2)

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

flextable() %>%

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

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

autofit(

}

# What does JumpeR do?

JumpeR is very similar to SwimmeR. They both mostly serve to convert results from human readable documents to machine & human readable data frames in the context of the R programming environment.

# Examples

## A Running Race

Here’s an example, reading in the 2019 Ivy League Championships and looking at the finals of the Women’s 200M Dash

df <- tf\_parse( read\_results("<http://www.leonetiming.com/2020/Indoor/IvyLeague/Results.htm>")

df %>%

filter(Event == "Women 200 Meter Dash") %>% group\_by(Name, Team) %>% # to remove prelims slice(2) %>% # to remove prelims arrange(Place) %>% # arrange by Place flextable\_style(

Place Name Age Team Finals\_Result Tiebreaker DQ Event

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 1 Katina Martin | SO Harvard | 24.05 |  | 0 | Women 200 Meter Dash |
| 2 Olivia Okoli | JR Harvard | 24.44 |  | 0 | Women 200 Meter Dash |
| 3 Cecil Ene | SR Penn | 24.52 | 24.511 | 0 | Women 200 Meter Dash |
| 4 Elena Brown-Soler | SR Penn | 24.52 | 24.520 | 0 | Women 200 Meter Dash |

Place Name Age Team Finals\_Result Tiebreaker DQ Event

1. Katie DiFrancesco JR Princeton 24.53 0 Women 200 Meter Dash
2. Libby McMahon SO Yale 25.12 0 Women 200 Meter Dash
3. Isabella Hilditch SO Princeton 40.06 0 Women 200 Meter Dash

Kennedy Waite FR Brown DNF 1 Women 200 Meter Dash

## Discus, with Flights

But wait, there’s more! Field events, like jumping and throwing, allow athletes to try several times, with each try called a “flight”. Flights can be captured as well. Here’s the Men’s Discus from the 2019 Virginia Grand Prix

df <- tf\_parse( read\_results("https://[www.flashresults.com/2019\_Meets/Outdoor/04-27\_](http://www.flashresults.com/2019_Meets/Outdoor/04-27_)

VirginiaGrandPrix/038-1.pdf"), flights = TRUE

df %>%

flextable\_style(

Place Name Age Team Finals\_Result DQ Event Flight\_1 Flight\_2 Flight\_3 Flight\_4 Flight\_5 Flight\_6

Nicholas

1

EDWARDS

FR HAMPTON 49.86m 0

Men Discus

X 47.11 45.99 47.28 X 49.86

Michael

2

ALBERT

JR APP STATE 48.30m 0

Men Discus

48.30 47.16 44.96 X 45.85 X

Joshua

3

HUNTER

SO HAMPTON 47.43m 0

Men Discus

31.94 X 46.54 X 47.43 X

1. Peter KENN SR APP STATE 46.14m 0

Men Discus

X 42.83 46.14 44.26 43.80 44.66

1. Asher PRINCE FR CHARLOTTE 45.98m 0

Men Discus

X 45.98 44.62 X X X

1. Sasha DAJIA SR CHARLOTTE 44.40m 0

Men Discus

X 44.40 44.19 X 44.08 42.04

1. Britton MANN SR HIGH POINT 42.07m 0

Men Discus

X 38.31 X 40.49 X 42.07

Gabriel

8

STAINBACK

SO HIGH POINT 39.37m 0

Men Discus

38.53 36.94 39.37 – – –

Place Name Age Team Finals\_Result DQ Event Flight\_1 Flight\_2 Flight\_3 Flight\_4 Flight\_5 Flight\_6

FOUL

Kysheen MYRICK

SO LIBERTY FOUL 1

Men Discus

X X X

FOUL Tyson JONES FR

VIRGINIA TECH

FOUL 1

Men Discus

X X X

## Pole Vault, with Flights and Attempts

JumpeR can even capture attempts for vertical jumping events, like in these Women’s Pole Vault results from the 2019 Texas A&M Invite. These results do get quite wide, so here they’re cut off at Flight 2.

df <- tf\_parse(

read\_results("https://[www.flashresults.com/2019\_Meets/Outdoor/04-12\_TamuInvite/014-1.pdf](http://www.flashresults.com/2019_Meets/Outdoor/04-12_TamuInvite/014-1.pdf)"), flights = TRUE

flight\_attempts = TRUE

df %>%

select(Place:Flight\_2\_Attempts) %>% flextable\_style(

Place Name Age Team Finals\_Result DQ Event Flight\_1 Flight\_1\_Attempts Flight\_2 Flight\_2\_Attempts

Caroline

1

BELLOWS

SR UTSA 3.88m 0

Women Pole Vault

3.28 — 3.43 O

Myka

2

STEINBEISSER

ARIZONA

FR

STATE

3.73m 0

Women Pole Vault

3.28 — 3.43 O

Tommi

3

HINTNAUS

ARIZONA

SO

STATE

3.73m 0

Women Pole Vault

3.28 — 3.43 —

4 Erika WILLIS FR

AIR FORCE

3.58m 0

Women Pole Vault

3.28 — 3.43 O

Kylie

5

SWIEKATOWSKI

JR RICE 3.58m 0

Women Pole Vault

3.28 — 3.43 XO

Cameron

6

BOEDEKER

SAM

JR HOUSTON 3.58m 0

ST.

Women Pole Vault

3.28 — 3.43 O

6 Kendahl SHUE JR TCU 3.58m 0

Women Pole Vault

3.28 — 3.43 —

Place Name Age Team Finals\_Result DQ Event Flight\_1 Flight\_1\_Attempts Flight\_2 Flight\_2\_Attempts

Corey

8

FRIEDENBACH

AIR

FR

FORCE

Women

Tysen

9

TOWNSEND

FR TCU

10 Lauren LABAY JR

SAM

HOUSTON 3.43m 0

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 3.58m | 0 | Pole Vault  Women | 3.28 | — | 3.43 | O |
| 3.58m | 0 | Pole  Vault | 3.28 | — | 3.43 | XXO |

ST.

Women Pole Vault

3.28 — 3.43 O

Margaret

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| SR | SAM  HOUSTON 3.43m | 0 | Women  Pole | 3.28 | — | 3.43 | O |
|  | ST. |  | Vault |  |  |  |  |
|  |  |  | Women |  |  |  |  |
| FR RICE 3.43m | | 0 | Pole Vault | 3.28 | — | 3.43 | XXO |
| AIR  FR 3.43m | | 0 | Women Pole | 3.28 | O | 3.43 | XXO |
|  | |  | Vault |  |  |  |  |
| FR UTSA DNS | | 0 | Women Pole | 3.28 |  | 3.43 |  |
|  | |  | Vault |  |  |  |  |
| TEXAS  SO NH | | 1 | Women Pole | 3.28 | — | 3.43 | XXX |
|  | |  | Vault |  |  |  |  |
| AIR  SR NH | | 1 | Women Pole | 3.28 | — | 3.43 | — |
|  | |  | Vault |  |  |  |  |
| AIR  SR NH | | 1 | Women Pole | 3.28 | — | 3.43 | — |
|  | |  | Vault |  |  |  |  |

10

LASSALLE

Emily

12

HARRISON

12

DNS

NH

Frankie PORAMBO

Alexandria GRAY

Hannah SEARBY

FORCE

A&M

NH Jerni SELF

FORCE

Kathryn

NH

TOMCZAK

FORCE

## Pole Vault Long Format

These results do get quite wide, but don’t worry. Switching to longer is easy as with

JumpeR::attempts\_split\_long.

df <- tf\_parse(

read\_results("https://[www.flashresults.com/2019\_Meets/Outdoor/04-12\_TamuInvite/014-1.pdf](http://www.flashresults.com/2019_Meets/Outdoor/04-12_TamuInvite/014-1.pdf)"), flights = TRUE

flight\_attempts = TRUE

df %>%

attempts\_split\_long() %>%

filter(Place == 1) %>% # only first place athlete

select(Place, Name, Age, Team, Finals\_Result, Event, Bar\_Height, Attempt, Result) %>%

flextable\_style(

Place Name Age Team Finals\_Result Event Bar\_Height Attempt Result

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 1 | Caroline BELLOWS SR | UTSA 3.88m | Women Pole Vault 3.28 | 1 | – |
| 1 | Caroline BELLOWS SR | UTSA 3.88m | Women Pole Vault 3.28 | 2 | – |
| 1 | Caroline BELLOWS SR | UTSA 3.88m | Women Pole Vault 3.28 | 3 | – |
| 1 | Caroline BELLOWS SR | UTSA 3.88m | Women Pole Vault 3.43 | 1 | O |
| 1 | Caroline BELLOWS SR | UTSA 3.88m | Women Pole Vault 3.58 | 1 | X |
| 1 | Caroline BELLOWS SR | UTSA 3.88m | Women Pole Vault 3.58 | 2 | O |
| 1 | Caroline BELLOWS SR | UTSA 3.88m | Women Pole Vault 3.73 | 1 | X |
| 1 | Caroline BELLOWS SR | UTSA 3.88m | Women Pole Vault 3.73 | 2 | O |
| 1 | Caroline BELLOWS SR | UTSA 3.88m | Women Pole Vault 3.88 | 1 | X |
| 1 | Caroline BELLOWS SR | UTSA 3.88m | Women Pole Vault 3.88 | 2 | O |
| 1 | Caroline BELLOWS SR | UTSA 3.88m | Women Pole Vault 4.03 | 1 | X |
| 1 | Caroline BELLOWS SR | UTSA 3.88m | Women Pole Vault 4.03 | 2 | X |
| 1 | Caroline BELLOWS SR | UTSA 3.88m | Women Pole Vault 4.03 | 3 | X |

## Relay Athletes

Going back to those Ivy League results, we can pull out the names relay athletes for each relay.

df <- tf\_parse( read\_results("<http://www.leonetiming.com/2020/Indoor/IvyLeague/Results.htm>"), relay\_athletes = TRUE

df %>%

filter(Event == "Men 4x400 Meter Relay") %>% select(-Tiebreaker, -Name) %>% flextable\_style(

Place Age Team Finals\_Result DQ Event Relay\_Athlete\_1 Relay\_Athlete\_2 Relay\_Athlete\_3 Relay\_Athlete\_4

Place Age Team Finals\_Result DQ Event Relay\_Athlete\_1 Relay\_Athlete\_2 Relay\_Athlete\_3 Relay\_Athlete\_4

1 Harvard 3:13.85 0

Men 4×400

Meter Relay

Aaron Shirley Gregory Lapit Charles Lego

Jovahn Williamson

2 Penn 3:15.55 0

Men 4×400

Meter Relay

Robbie Ruppel Anthony Okolo Emerson Douds

Antaures Jackson

3 Yale 3:16.60 0

Men 4×400

Meter Relay

Christopher Colbert

Juma Sei Phil Zuccaro Marcus Woods

4 Cornell 3:17.61 0

Men 4×400

Meter Relay

Christian Martin Myles Solan

Malick Diomande

Tien Henderson

5 Dartmouth 3:17.66 0

Men 4×400

Meter Relay

Mathieu Farber Charlie Wade Julian Martelly Max Frye

6 Columbia 3:19.42 0

Men 4×400

Meter Relay

Chris Balthazar Jahi Hernandez Brodie Holmes Vasilis Kopanas

7 Princeton 3:20.61 0

Men 4×400

Meter Relay

Gregory Sholars Klaudio Gjetja

Anderson Dimon

Michael Phillippy

8 Brown 3:25.72 0

Men 4×400

Meter Relay

Sergey Gorban Austin Reynolds Kevin Boyce Tim McDonough

## Formating Results

Track and field results are of two forms. Times, as “MM:SS.HH”, and lengths/heights, often as “X.XXm”. JumpeR has math\_format for converting these result strings into numerics, which is useful when doing comparisons and plotting. Here’s the men’s pole vault at the USA T&F 2019 Championships .

df <- tf\_parse(

read\_results("https://[www.flashresults.com/2019\_Meets/Outdoor/07-25\_USATF\_CIS/026-1.pdf](http://www.flashresults.com/2019_Meets/Outdoor/07-25_USATF_CIS/026-1.pdf)"))

df %>%

mutate(Finals\_Math = math\_format(Finals\_Result)) %>% # results to numerics

mutate(Name = factor(Name, unique(Name))) %>% # order names by order of finish ggplot(aes(x = Name, y = Finals\_Math)) +

geom\_col() + theme\_bw() +

theme(axis.text.x = element\_text( angle = 90

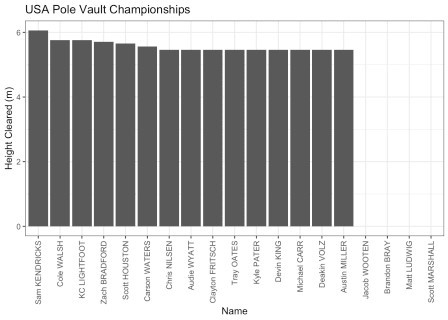
vjust = 0.5

hjust = 1

)) +

labs(y = "Height Cleared (m)"

title = "USA Pole Vault Championships"



One can use math\_format on mixed format lists too. Times will be converted to seconds, meters will remain in meters, and standard units (feet, inches) will be converted to inches. Units however are not included, so be aware.

demo\_list <- c(

"1.23m", # a height/length in meters, output in meters

"5-06.45", # a height/length in standard, output in inches "10:34.34", # a time with minutes, output in seconds "9.45" # a time without minutes, output in seconds

)

math\_format(demo\_list

## [1] 1.23 66.45 634.34 9.45

# JumpeR Going Forward