

Got R? Catch 'em all!

Pluck the data from Bulbapedia

PokedexR

```
library(rvest)
library(magrittr)

bulbagarden <- "http://bulbapedia.bulbagarden.net/wiki/List
baseStats <-
  # Read HTML or XML
  xml2::read_html(
    x = bulbagarden
  ) %>% # This is a pipe operator from magrittr
  # Extract pieces out of HTML using css selectors
  rvest::html_node(
    # rvest recommends using 'Selector Gadget'
    css = "div table"
  ) %>%
  # Parse an html table into a data frame
  rvest::html_table()
```

All the data

```
head(baseStats)
```

##	#		Pokémon	HP	Attack	Defense	Speed	Special	Total	Average	
##	1	1	NA	Bulbasaur	45	49	49	45	65	253	50.6
##	2	2	NA	Ivysaur	60	62	63	60	80	325	65.0
##	3	3	NA	Venusaur	80	82	83	80	100	425	85.0
##	4	4	NA	Charmander	39	52	43	65	50	249	49.8
##	5	5	NA	Charmeleon	58	64	58	80	65	325	65.0
##	6	6	NA	Charizard	78	84	78	100	85	425	85.0

SelectorGadget (like a silph scope but for web pages)

Invest recommends SelectorGadget is a chrome extension for CSS selector generation.

It “Makes the Invisible Plain to See!” by exposing which parts of the html correspond to which bits of the user facing webpage.

The screenshot shows the SelectorGadget interface overlaid on a webpage. The webpage content includes a heading "List of Pokémon by base stats" and a table of Pokémon stats. The SelectorGadget interface shows the selected CSS selector: `.jquery-tablesorter , .headerSort , td~ td+ td a`. The interface also includes buttons for "Clear (161)", "Toggle Position", "XPath", and a search icon.

stats, able to be changed at will to list them in order of any one of the five stat categories, as well as the average of all and then

Pokédex number.

List of Pokémon by base stats

h2 span

#	Pokémon	HP	Attack	Defense	Speed	Special	Total	Average
001	Bulbasaur	45	49	49	45	65	253	50.6
003	Venusaur	80	82	83	80	100	425	85

.jquery-tablesorter , .headerSort , td~ td+ td a

Clear (161) Toggle Position XPath ? X

Figure 1:SelectorGadget

However, I had to fudge it a bit, as it didn't pick up the table properly.

Muky Data

```
library(data.table)

baseStats <- data.table::as.data.table(baseStats)

# Remove second col with only "NA"
baseStats[, "" := NULL]
# Rename cols 1 and 2 to something workable
data.table::setnames(baseStats,
                      1:2,
                      c("DexNo", "Pokemon"))
```

A better Pokeball

Master Ballin'

```
library(stringr)

baseStats[, imgURL := read_html(
  x = bulbagarden# set x to be bulbagarden url
) %>%
rvest::html_nodes(
  # css to identify all strings for pokemon image urls
  css = "#mw-content-text img"
) %>% # split string at point
stringr::str_split_fixed(
  "src=\"",
  n = 2
) %>% # use second part of string
  .[,2] %>% # split string at point again
stringr::str_split_fixed(
  "\" width=",
  n = 2) %>% # use first part of string
  .[,1]]
```


Congratulations, you caught all 151 image urls!

```
head(baseStats)
```

```
##      DexNo   Pokemon HP Attack Defense Speed Special Total Ave
## 1:      1 Bulbasaur 45    49    49    45    65   253   50.6
## 2:      2 Ivysaur 60    62    63    60    80   325   65.0
## 3:      3 Venusaur 80    82    83    80   100   425   85.0
## 4:      4 Charmander 39    52    43    65    50   249   49.8
## 5:      5 Charmeleon 58    64    58    80    65   325   65.0
## 6:      6 Charizard 78    84    78   100    85   425   85.0
##                                     imgURL
## 1: http://cdn.bulbagarden.net/upload/e/ec/001MS.png
## 2: http://cdn.bulbagarden.net/upload/6/6b/002MS.png
## 3: http://cdn.bulbagarden.net/upload/d/df/003MS.png
## 4: http://cdn.bulbagarden.net/upload/b/bb/004MS.png
## 5: http://cdn.bulbagarden.net/upload/d/dc/005MS.png
## 6: http://cdn.bulbagarden.net/upload/0/01/006MS.png
```

Your Pokemon have been moved to "Someones PC"!

```
dir.create("./someonesPC/")

# Use for loop as data.table only
# returns last item if this is run within DT function
for(pkmn in baseStats[, DexNo]){
  download.file(baseStats[pkmn == DexNo,
                        imgURL],
                destfile = paste0("./someonesPC/",
                                    pkmn, ".png"),
                # Makes this work for windows
                mode = "wb")
}
```

DaveRGP checked “Someones PC”!

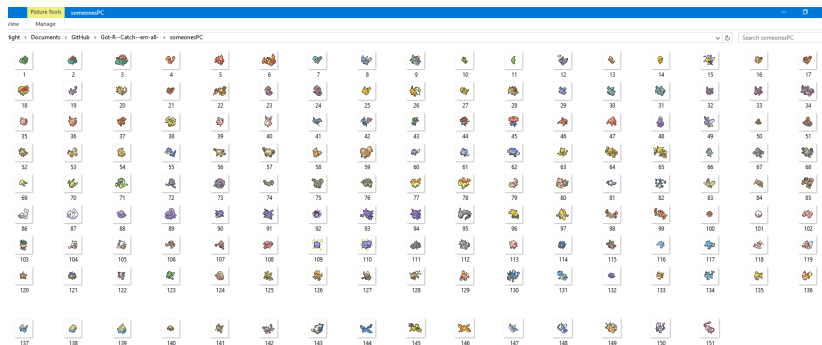


Figure 2:someones PC

I'm gonna be...

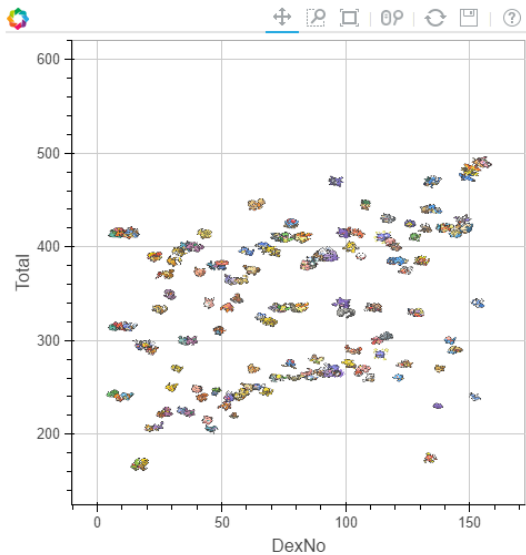
...the very best...

```
library(rbokeh)

P <- figure() %>%
  # layer to get .png from url as points
  ly_image_url(
    data = baseStats,
    x = DexNo,
    y = Total,
    w = 10,
    h = 20,
    image_url = imgURL
  )
```

... like no one ever was!

P



Trainer Card