Rectangling Lab

Sha Lu

6/14/2021

library(rvest)  
library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

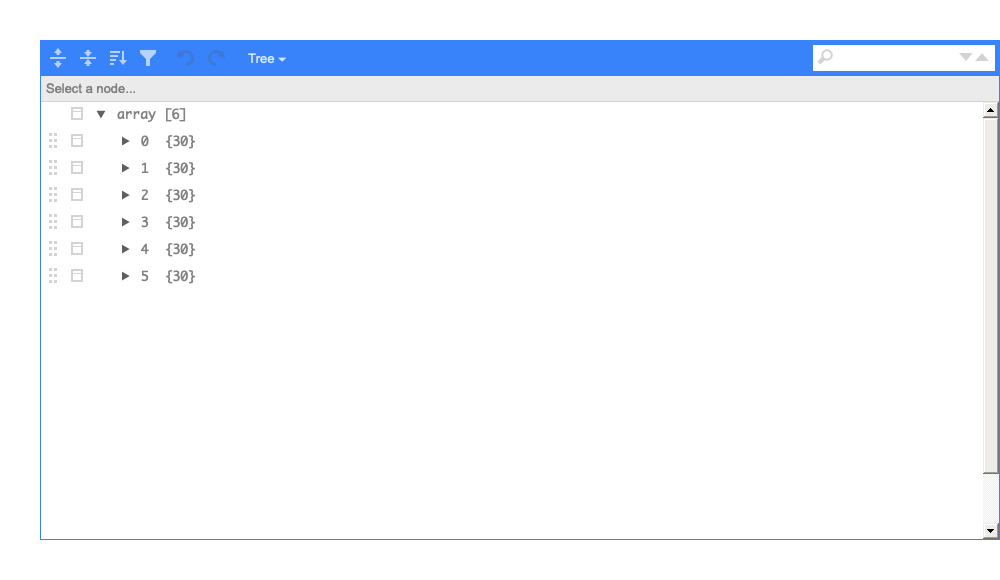
library(tidyverse)

## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.0 ──

## ✓ ggplot2 3.3.3 ✓ purrr 0.3.4  
## ✓ tibble 3.0.5 ✓ stringr 1.4.0  
## ✓ tidyr 1.1.2 ✓ forcats 0.5.1  
## ✓ readr 1.4.0

## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## x dplyr::filter() masks stats::filter()  
## x readr::guess\_encoding() masks rvest::guess\_encoding()  
## x dplyr::lag() masks stats::lag()

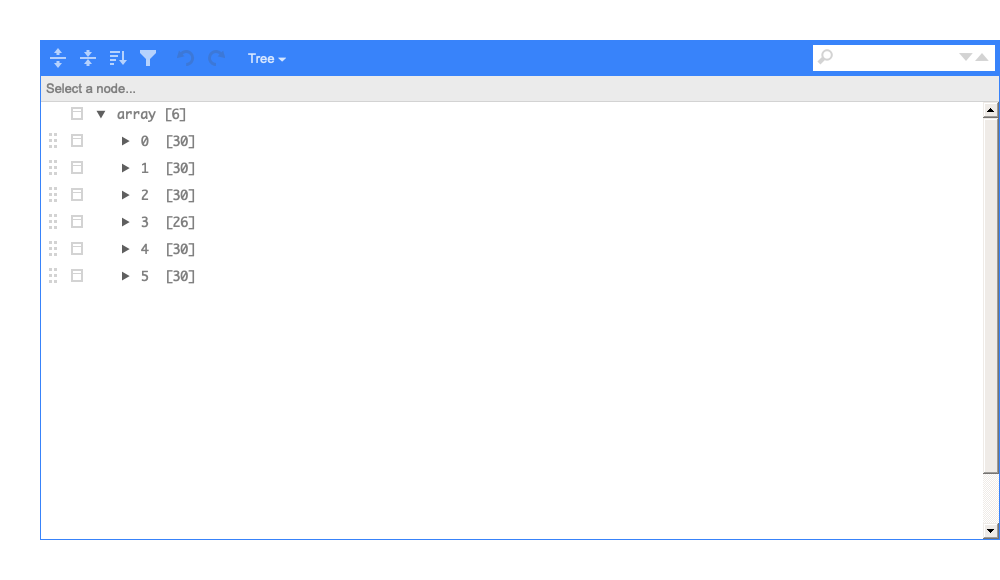
library(repurrrsive)  
listviewer::jsonedit(gh\_users)

 a) Explain or describe nested Lists.

## Nested list means there are lists inside a list.

1. Open and explore data for Git Hub repos by running the following code gh\_repos. Comment on the structure and arrangement of the data (Two or three sentences)

listviewer::jsonedit(gh\_repos)



## This is an array with 6 nested lists inside. Every list except list 3 has 30 sub-lists, list 3 has 26 sub-lists. Under the sub-lists, there are 68 elements or objects sepreatly.

1. In two or three sentences, explain what happens to the data when you run the code

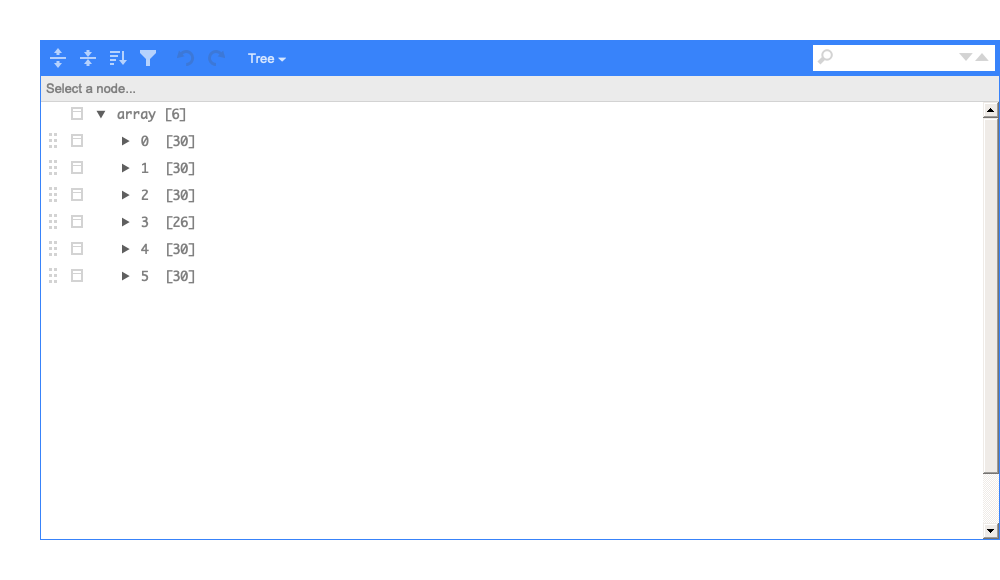
repos <- tibble(repo = gh\_repos)  
repos

## # A tibble: 6 x 1  
## repo   
## <list>   
## 1 <list [30]>  
## 2 <list [30]>  
## 3 <list [30]>  
## 4 <list [26]>  
## 5 <list [30]>  
## 6 <list [30]>

## This code organizes the data into a tibble. Whole repo has 6 lists. Every list has 30 sub lists.

1. Describe the data representation when the listserver tool is applied. And go into detail explaining why we have a nested List configuration. What do the numbers 30 and 68 represent ? (6 or 7 sentences)

listviewer::jsonedit(gh\_repos)



## This array has 6 lists. Under every list, there are 30 sub-lists. Except list 3 has only 26 sub-list. Every sub-list is a user's information, it has 68 elements.

tibble(repo = gh\_repos)%>%  
 unnest\_auto(repo)%>%  
 unnest\_auto(repo) -> new

## Using `unnest\_longer(repo)`; no element has names

## Using `unnest\_wider(repo)`; elements have 68 names in common

new

## # A tibble: 176 x 67  
## id name full\_name owner private html\_url description fork url   
## <int> <chr> <chr> <lis> <lgl> <chr> <chr> <lgl> <chr>  
## 1 6.12e7 after gaborcsa… <nam… FALSE https:/… Run Code i… FALSE http…  
## 2 4.05e7 argu… gaborcsa… <nam… FALSE https:/… Declarativ… FALSE http…  
## 3 3.64e7 ask gaborcsa… <nam… FALSE https:/… Friendly C… FALSE http…  
## 4 3.49e7 base… gaborcsa… <nam… FALSE https:/… Do we get … FALSE http…  
## 5 6.16e7 cite… gaborcsa… <nam… FALSE https:/… Test R pac… TRUE http…  
## 6 3.39e7 clis… gaborcsa… <nam… FALSE https:/… Unicode sy… FALSE http…  
## 7 3.72e7 cmak… gaborcsa… <nam… FALSE https:/… port of cm… TRUE http…  
## 8 6.80e7 cmark gaborcsa… <nam… FALSE https:/… CommonMark… TRUE http…  
## 9 6.32e7 cond… gaborcsa… <nam… FALSE https:/… <NA> TRUE http…  
## 10 2.43e7 cray… gaborcsa… <nam… FALSE https:/… R package … FALSE http…  
## # … with 166 more rows, and 58 more variables: forks\_url <chr>, keys\_url <chr>,  
## # collaborators\_url <chr>, teams\_url <chr>, hooks\_url <chr>,  
## # issue\_events\_url <chr>, events\_url <chr>, assignees\_url <chr>,  
## # branches\_url <chr>, tags\_url <chr>, blobs\_url <chr>, git\_tags\_url <chr>,  
## # git\_refs\_url <chr>, trees\_url <chr>, statuses\_url <chr>,  
## # languages\_url <chr>, stargazers\_url <chr>, contributors\_url <chr>,  
## # subscribers\_url <chr>, subscription\_url <chr>, commits\_url <chr>,  
## # git\_commits\_url <chr>, comments\_url <chr>, issue\_comment\_url <chr>,  
## # contents\_url <chr>, compare\_url <chr>, merges\_url <chr>, archive\_url <chr>,  
## # downloads\_url <chr>, issues\_url <chr>, pulls\_url <chr>,  
## # milestones\_url <chr>, notifications\_url <chr>, labels\_url <chr>,  
## # releases\_url <chr>, deployments\_url <chr>, created\_at <chr>,  
## # updated\_at <chr>, pushed\_at <chr>, git\_url <chr>, ssh\_url <chr>,  
## # clone\_url <chr>, svn\_url <chr>, size <int>, stargazers\_count <int>,  
## # watchers\_count <int>, language <chr>, has\_issues <lgl>,  
## # has\_downloads <lgl>, has\_wiki <lgl>, has\_pages <lgl>, forks\_count <int>,  
## # open\_issues\_count <int>, forks <int>, open\_issues <int>, watchers <int>,  
## # default\_branch <chr>, homepage <chr>