STAT 613 - Lab for Rectangling

Rajeev Agrawal

6/14/2021

## 1)

### a) Explain or describe nested Lists. (Three or four sentences)

A list can contain sublists, which in turn can contain sublists themselves, and so on. This is known as nested list. Sometimes the lists are contained in another list but we want to access the nested list’s elements. Since these elements are part of a list that cannot be directly accessed, first we need to access the broader list and then the list that contains the element to reach the actual element.

### b) 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)

It is an example of a nested list, that is, there are many lists contained inside the gh\_repos, which itself is a list.

gh\_repos

### c) In two or three sentences, explain what happens to the data when you run the code;

The code converts gh\_repos into a tibble (a rectangular data frame) of lists that were originally contained inside gh\_repos. There are 6 repos, each containing a list of 30 lists except for 4th repo, which contains only 26 lists.

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]>

### d) 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 lets us look at the list data for each Github repository in gh\_repos. We already know that gh\_repos is a nested list, that it a list containing lists, which in turn also contains lists, which is why we have a nested list configuration in the listviewer tool.

30 (or 26 for the 4th repo) represents the number of lists in each of the repos in gh\_repo. 68 on the other hand, represents the number of lists in each of the 30 sublists inside the 6 repo lists in gh\_repo.

listviewer::jsonedit(gh\_repos)

### e) Use and show R unesting code to produce the tibble below.

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

## # 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 gaborcsar~ <nam~ FALSE https://~ Run Code in ~ FALSE https:~  
## 2 4.05e7 argufy gaborcsar~ <nam~ FALSE https://~ Declarative ~ FALSE https:~  
## 3 3.64e7 ask gaborcsar~ <nam~ FALSE https://~ Friendly CLI~ FALSE https:~  
## 4 3.49e7 basei~ gaborcsar~ <nam~ FALSE https://~ Do we get wa~ FALSE https:~  
## 5 6.16e7 citest gaborcsar~ <nam~ FALSE https://~ Test R packa~ TRUE https:~  
## 6 3.39e7 clisy~ gaborcsar~ <nam~ FALSE https://~ Unicode symb~ FALSE https:~  
## 7 3.72e7 cmaker gaborcsar~ <nam~ FALSE https://~ port of cmak~ TRUE https:~  
## 8 6.80e7 cmark gaborcsar~ <nam~ FALSE https://~ CommonMark p~ TRUE https:~  
## 9 6.32e7 condi~ gaborcsar~ <nam~ FALSE https://~ <NA> TRUE https:~  
## 10 2.43e7 crayon gaborcsar~ <nam~ FALSE https://~ R package fo~ FALSE https:~  
## # ... 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>