Rectangling\_Lab

Madeleine Augostini

6/14/2021

library(rvest)   
library(repurrrsive)  
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 3.0.4 ✓ stringr 1.4.0  
## ✓ tidyr 1.1.2 ✓ forcats 0.5.0  
## ✓ 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()

#listviewer::jsonedit(gh\_users) #commented out to reduce output

1.)

1. Explain or describe nested Lists A nested list is when there is a list within a list. A list is a vector with objects of different types, e.g., integer, character, etc., but this also includes when one or more elements are other lists. It can be difficult to read and interpret data in a nested list and we often will unnest a list before manipulating it.

#gh\_repos #commented out to reduce output

At first glance, this data is very messy and difficult to interpret. Much of the information appears to be nested, perhaps more than one layer (a list within a list within another list), and some of the data appears to be api URLs for various information.

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

This format now shows us the main list consists of six other lists. Each of the six lists contain either 26 or 30 elements, however at this point, we are not able to see what the lists contain.

#listviewer::jsonedit(gh\_repos)

Using the listserver tool allows us to expand/collapse each to the lists. We can see the lists are nested even further, and when expanded all the way, we are able to see the contents of the objects. The first nested list contains 30 other lists, each which contain 68 objects. When we open these lists, we can see that the 68 objects include information such as userID, descriptions, api urls, and other identifying information/repositorites. The objects contain information regarding one user. This is an especially useful way to view the data as a whole since it is interactive and fits on one page. We can use this information to help come up with a method for further manipulating the data later.

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>