Homework 4

Sneha Karanjai

9/18/2022

Introduction

The goal of this homework is to get practice reading in raw data from different sources. Some files are available in the homework link, others via a URL, and others may be connected to using a package of some type.

Part 1

1. If your working directory is `myfolder/homework/`, what path would you specify to get the file located

at `myfolder/MyData.csv`?

To get the data MyData.csv, the path we specify is `myfolder/MyData.csv`

2. What are the major benefits of using R projects? Should you be using an R project for each homework assignment (or at least for the course)??

Projects help us in creating reproducible code. Packages help us in collaboration through Version Control in Github. The way packages help in collaboration is that if we build individual R programs and then share it across to other people then our working directory will not match theirs. Hence, they would have to change the paths for individual file imports. Projects help in making sure that this problem does not exist. To summarise, Projects help in : - Collaborating - Reproducibility - Easier Imports

Instead of creating a new project for each homework, it is better to have a single project for the entire course.

3. What is git and what is github?

Git is an open-source, version control tool created in 2005 by developers working on the Linux operating system; GitHub is a company founded in 2008 that makes tools which integrate with git. You do not need GitHub to use git, but you cannot use GitHub without using git.

Part 2

The purpose of this part is to read in delimited dataset in R. The source of the datasets is UCI Machine Learning Repository.

Importing necessary libraries

```
library(readr)
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
```

Glass Data

1. Read this data into R using functions from the tidyverse. Notice that the data doesn't include column

names - add those (in a manner of your choosing). Print out the tibble (just call the object name).

This dataset is a comma delimited or a .csv file. This means that the value of each column is separated by a comma. To read in the data we will use the `read_csv()` function from the tidyverse's `readr` package. Reading the data into R environment involves adding the column names manually as the dataframe in its original format does not include the header row. Print it out as a `tibble`.

```
glass_data <- read_csv("https://www4.stat.ncsu.edu/~online/datasets/glass.data",
                       col_names = c("Id", "RI", "Na", "Mg", "Al", "Si", "K", "Ca", "Ba", "Fe", "Type_O
```

```
## Rows: 214 Columns: 11

## -- Column specification --------------------------------------------------------
## Delimiter: ","
## dbl (11): Id, RI, Na, Mg, Al, Si, K, Ca, Ba, Fe, Type_Of_Glass

##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
glass_data
```

```
## # A tibble: 214 x 11
##       Id    RI    Na    Mg    Al    Si     K    Ca    Ba    Fe Type_Of_Glass
##    <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>         <dbl>
##  1     1  1.52  13.6  4.49  1.1   71.8  0.06  8.75     0  0                1
##  2     2  1.52  13.9  3.6   1.36  72.7  0.48  7.83     0  0                1
##  3     3  1.52  13.5  3.55  1.54  73.0  0.39  7.78     0  0                1
##  4     4  1.52  13.2  3.69  1.29  72.6  0.57  8.22     0  0                1
```

```
##  5     5  1.52  13.3  3.62  1.24  73.1  0.55  8.07     0  0                1
##  6     6  1.52  12.8  3.61  1.62  73.0  0.64  8.07     0  0.26             1
##  7     7  1.52  13.3  3.6   1.14  73.1  0.58  8.17     0  0                1
##  8     8  1.52  13.2  3.61  1.05  73.2  0.57  8.24     0  0                1
##  9     9  1.52  14.0  3.58  1.37  72.1  0.56  8.3      0  0                1
## 10    10  1.52  13    3.6   1.36  73.0  0.57  8.4      0  0.11             1
## # ... with 204 more rows
```

2. Overwrite the `Type_of_glass` variable by creating a factor there instead. Use the variable descriptions above to give meaningful factor levels.

We see that the `Type_Of_Glass` column is read in as a double because of the values of the columns being from 1-7. We need to overwrite this to a factor variable. Factor in R is a variable used to categorize and store the data, having a limited number of different values. We use the `factor()` function from the base package.

```
glass_data$Type_Of_Glass <- factor(glass_data$Type_Of_Glass)
glass_data
```

```
## # A tibble: 214 x 11
##       Id    RI    Na    Mg    Al    Si     K    Ca    Ba    Fe Type_Of_Glass
##    <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <fct>
##  1     1  1.52  13.6  4.49  1.1   71.8  0.06  8.75     0  0    1
##  2     2  1.52  13.9  3.6   1.36  72.7  0.48  7.83     0  0    1
##  3     3  1.52  13.5  3.55  1.54  73.0  0.39  7.78     0  0    1
##  4     4  1.52  13.2  3.69  1.29  72.6  0.57  8.22     0  0    1
##  5     5  1.52  13.3  3.62  1.24  73.1  0.55  8.07     0  0    1
##  6     6  1.52  12.8  3.61  1.62  73.0  0.64  8.07     0  0.26 1
##  7     7  1.52  13.3  3.6   1.14  73.1  0.58  8.17     0  0    1
##  8     8  1.52  13.2  3.61  1.05  73.2  0.57  8.24     0  0    1
##  9     9  1.52  14.0  3.58  1.37  72.1  0.56  8.3      0  0    1
## 10    10  1.52  13    3.6   1.36  73.0  0.57  8.4      0  0.11 1
## # ... with 204 more rows
```

Now that we have converted the data type of `Type_Of_Glass`, it is time for us re-label the factors. The labels need to be factored :

- 1 building_windows_float_processed
- 2 building_windows_non_float_processed
- 3 vehicle_windows_float_processed
- 4 vehicle_windows_non_float_processed (none in this database)
- 5 containers
- 6 tableware
- 7 headlamps

For this we use the `recode()` function from the `dplyr` package.

```
glass_data$Type_Of_Glass <- recode(glass_data$Type_Of_Glass,
                                   "1" = "building_windows_float_processed",
                                   "2" = "building_windows_non_float_processed",
                                   "3" = "vehicle_windows_float_processed",
                                   "4" = "vehicle_windows_non_float_processed",
```

```
                                          "5" = "containers",
                                          "6" = "tableware",
                                          "7" = "headlamps")
glass_data
```

```
## # A tibble: 214 x 11
##       Id    RI    Na    Mg    Al    Si     K    Ca    Ba    Fe Type_Of_Glass
##    <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <fct>
##  1     1  1.52  13.6  4.49  1.1   71.8  0.06  8.75     0  0    building_windows~
##  2     2  1.52  13.9  3.6   1.36  72.7  0.48  7.83     0  0    building_windows~
##  3     3  1.52  13.5  3.55  1.54  73.0  0.39  7.78     0  0    building_windows~
##  4     4  1.52  13.2  3.69  1.29  72.6  0.57  8.22     0  0    building_windows~
##  5     5  1.52  13.3  3.62  1.24  73.1  0.55  8.07     0  0    building_windows~
##  6     6  1.52  12.8  3.61  1.62  73.0  0.64  8.07     0  0.26 building_windows~
##  7     7  1.52  13.3  3.6   1.14  73.1  0.58  8.17     0  0    building_windows~
##  8     8  1.52  13.2  3.61  1.05  73.2  0.57  8.24     0  0    building_windows~
##  9     9  1.52  14.0  3.58  1.37  72.1  0.56  8.3      0  0    building_windows~
## 10    10  1.52  13    3.6   1.36  73.0  0.57  8.4      0  0.11 building_windows~
## # ... with 204 more rows
```

3. Print the data frame with only observations where the `Fe` variable is less than 0.2 and the `Type of Glass`is either tableware or headlamp.

Success! Now we filter the dataset using the `filter()` function to display records with `Fe` value less than 0.2 and `Type_Of_Glass` value either tableware or headlamps.

```
glass_data %>%
  filter((Fe < 0.2) & Type_Of_Glass %in% c("tableware", "headlamps"))
```

```
## # A tibble: 38 x 11
##       Id    RI    Na    Mg    Al    Si     K    Ca    Ba    Fe Type_Of_Glass
##    <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <fct>
##  1   177  1.52  14    2.39  1.56  72.4  0     9.57  0        0  tableware
##  2   178  1.52  13.8  2.41  1.19  72.8  0     9.77  0        0  tableware
##  3   179  1.52  14.5  2.24  1.62  72.4  0     9.26  0        0  tableware
##  4   180  1.52  14.1  2.19  1.66  72.7  0     9.32  0        0  tableware
##  5   181  1.51  14.4  1.74  1.54  74.6  0     7.59  0        0  tableware
##  6   182  1.52  15.0  0.78  1.74  72.5  0     9.95  0        0  tableware
##  7   183  1.52  14.2  0     2.09  72.7  0    10.9   0        0  tableware
##  8   184  1.52  14.6  0     0.56  73.5  0    11.2   0        0  tableware
##  9   185  1.51  17.4  0     0.34  75.4  0     6.65  0        0  tableware
## 10   186  1.51  13.7  3.2   1.81  72.8  1.76  5.43  1.19     0  headlamps
## # ... with 28 more rows
```

Yeast Data

1. Read this data into R using functions from the tidyverse. Notice that the data doesn't include column names - add those (in a manner of your choosing). Print out the tibble (just call the object name).

The first task is the read the data from the URL provided. We notice that the raw form of the dataset does not include column names so we manually put that in while reading the dataset based on the information

provided to us.

```
yeast_data <- read_table("https://www4.stat.ncsu.edu/~online/datasets/yeast.data",
                         col_names = c("seq_name", "mcg", "gvh", "alm", "mit", "erl", "pox", "vac", "nuc"
```

```
##
## -- Column specification --------------------------------------------------------
## cols(
##   seq_name = col_character(),
##   mcg = col_double(),
##   gvh = col_double(),
##   alm = col_double(),
##   mit = col_double(),
##   erl = col_double(),
##   pox = col_double(),
##   vac = col_double(),
##   nuc = col_double(),
##   class = col_character()
## )
```

```
yeast_data
```

```
## # A tibble: 1,484 x 10
##    seq_name     mcg   gvh   alm   mit   erl   pox   vac   nuc class
##    <chr>      <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <chr>
##  1 ADT1_YEAST  0.58  0.61  0.47  0.13   0.5   0    0.48  0.22 MIT
##  2 ADT2_YEAST  0.43  0.67  0.48  0.27   0.5   0    0.53  0.22 MIT
##  3 ADT3_YEAST  0.64  0.62  0.49  0.15   0.5   0    0.53  0.22 MIT
##  4 AAR2_YEAST  0.58  0.44  0.57  0.13   0.5   0    0.54  0.22 NUC
##  5 AATM_YEAST  0.42  0.44  0.48  0.54   0.5   0    0.48  0.22 MIT
##  6 AATC_YEAST  0.51  0.4   0.56  0.17   0.5   0.5  0.49  0.22 CYT
##  7 ABC1_YEAST  0.5   0.54  0.48  0.65   0.5   0    0.53  0.22 MIT
##  8 BAF1_YEAST  0.48  0.45  0.59  0.2    0.5   0    0.58  0.34 NUC
##  9 ABF2_YEAST  0.55  0.5   0.66  0.36   0.5   0    0.49  0.22 MIT
## 10 ABP1_YEAST  0.4   0.39  0.6   0.15   0.5   0    0.58  0.3  CYT
## # ... with 1,474 more rows
```

2. Select only the class and mcg columns. Report the mean and standard deviation of the mcg value for each setting of the class variable

Now that we have the data in, it is time for data manipulation activities. We select class and mcg, we then group_by class and finally summarise the dataset to get the mean and standard deviation of mcg.

```
yeast_data %>%
  select(class, mcg) %>%
  group_by(class) %>%
  summarise(mean_mcg = mean(mcg), std_mcg = sd(mcg))
```

```
## # A tibble: 10 x 3
##    class mean_mcg std_mcg
##    <chr>    <dbl>   <dbl>
```

```
##  1 CYT      0.481  0.107
##  2 ERL      0.792  0.0653
##  3 EXC      0.735  0.111
##  4 ME1      0.789  0.0671
##  5 ME2      0.722  0.160
##  6 ME3      0.431  0.0989
##  7 MIT      0.521  0.0972
##  8 NUC      0.452  0.111
##  9 POX      0.521  0.133
## 10 VAC      0.548  0.141
```

Part 3 - Database

We will be working with an example SQLite database called chinook.

1. Download the chinook.db database. (If needed install and) load the DBI and RSQLite packages, and load the tidyverse package. Use dbConnect() to connect to the this local database.

First step is to download the chinook.db from Moodle. We then install and load `DBI` and `RSQLite` packages, alongwith the `tidyverse` package. Further, we use `dbConnect()` to connect to the local DB to reach out to the chinook.db.

```
#install.packages("DBI")
#install.packages("RSQLite")
library("DBI")
```

```
## Warning: package 'DBI' was built under R version 4.1.2
```

```
library("RSQLite")
```

```
## Warning: package 'RSQLite' was built under R version 4.1.2
```

```
library("dbplyr")
```

```
##
## Attaching package: 'dbplyr'

## The following objects are masked from 'package:dplyr':
##
##     ident, sql
```

```
con <- dbConnect(RSQLite::SQLite(), "chinook.db")
dbListTables(conn = con)
```

```
##  [1] "albums"          "artists"         "customers"       "employees"
##  [5] "genres"          "invoice_items"   "invoices"        "media_types"
##  [9] "playlist_track"  "playlists"       "sqlite_sequence" "sqlite_stat1"
## [13] "tracks"
```

2. Now print out the tables in the database using `dbListTables()`.

The `dbListTables()` function lists out all the tables in the given database.

```
dbListTables(conn = con)
```

```
##  [1] "albums"          "artists"         "customers"       "employees"
##  [5] "genres"          "invoice_items"   "invoices"        "media_types"
##  [9] "playlist_track"  "playlists"       "sqlite_sequence" "sqlite_stat1"
## [13] "tracks"
```

3. Use `dbGetQuery()` or `tbl()` to grab and print out the invoices table and the customers table.

We use `tbl()` which is a generic method that dispatches based on the first argument. We grab the invoices and customers table.

```
invoices <- tbl(con, "invoices")
invoices
```

```
## # Source:   table<invoices> [?? x 9]
## # Database: sqlite 3.39.3
## #   [/Users/snehakaranjai/Documents/Sem3/ST558/sneha-k.github.io/Homeworks/HW4/chinook.db]
##    InvoiceId CustomerId InvoiceDate    BillingAddress   BillingCity BillingState
##        <int>      <int> <chr>          <chr>            <chr>       <chr>
##  1         1          2 2009-01-01 00~ Theodor-Heuss-S~ Stuttgart   <NA>
##  2         2          4 2009-01-02 00~ Ullevålsveien 14 Oslo        <NA>
##  3         3          8 2009-01-03 00~ Grétrystraat 63  Brussels    <NA>
##  4         4         14 2009-01-06 00~ 8210 111 ST NW   Edmonton    AB
##  5         5         23 2009-01-11 00~ 69 Salem Street  Boston      MA
##  6         6         37 2009-01-19 00~ Berger Straße 10 Frankfurt   <NA>
##  7         7         38 2009-02-01 00~ Barbarossastraß~ Berlin      <NA>
##  8         8         40 2009-02-01 00~ 8, Rue Hanovre   Paris       <NA>
##  9         9         42 2009-02-02 00~ 9, Place Louis ~ Bordeaux    <NA>
## 10        10         46 2009-02-03 00~ 3 Chatham Street Dublin      Dublin
## # ... with more rows, and 3 more variables: BillingCountry <chr>,
## #   BillingPostalCode <chr>, Total <dbl>
```

```
customers <- tbl(con, "customers")
customers
```

```
## # Source:   table<customers> [?? x 13]
## # Database: sqlite 3.39.3
## #   [/Users/snehakaranjai/Documents/Sem3/ST558/sneha-k.github.io/Homeworks/HW4/chinook.db]
##    CustomerId FirstName LastName  Company Address City  State Country PostalCode
##         <int> <chr>     <chr>     <chr>   <chr>   <chr> <chr> <chr>   <chr>
##  1          1 Luís      Gonçalves Embrae~ Av. Br~ São ~ SP    Brazil  12227-000
##  2          2 Leonie    Köhler    <NA>    Theodo~ Stut~ <NA>  Germany 70174
##  3          3 François  Tremblay  <NA>    1498 r~ Mont~ QC    Canada  H2G 1A7
##  4          4 Bjørn     Hansen    <NA>    Ullevå~ Oslo  <NA>  Norway  0171
##  5          5 František Wichterl~ JetBra~ Klanov~ Prag~ <NA>  Czech ~ 14700
##  6          6 Helena    Holý      <NA>    Rilská~ Prag~ <NA>  Czech ~ 14300
##  7          7 Astrid    Gruber    <NA>    Rotent~ Vien~ <NA>  Austria 1010
```

```
##  8          8 Daan      Peeters   <NA>    Grétry~ Brus~ <NA>  Belgium 1000
##  9          9 Kara      Nielsen   <NA>    Sønder~ Cope~ <NA>  Denmark 1720
## 10         10 Eduardo   Martins   Woodst~ Rua Dr~ São ~ SP    Brazil  01007-010
## # ... with more rows, and 4 more variables: Phone <chr>, Fax <chr>,
## #   Email <chr>, SupportRepId <int>
```

4. Use an inner_join() to combine the two tables above by the CustomerID variable.

`inner_join()` is the mutating joins add columns from y to x, matching rows based on the keys that includes all rows in both x and y.

```
invoices %>%
  inner_join(customers, by = "CustomerId")
```

```
## # Source:   lazy query [?? x 21]
## # Database: sqlite 3.39.3
## #   [/Users/snehakaranjai/Documents/Sem3/ST558/sneha-k.github.io/Homeworks/HW4/chinook.db]
##    InvoiceId CustomerId InvoiceDate   BillingAddress   BillingCity  BillingState
##        <int>      <int> <chr>         <chr>            <chr>        <chr>
##  1        98          1 2010-03-11 0~ Av. Brigadeiro ~ São José do~ SP
##  2       121          1 2010-06-13 0~ Av. Brigadeiro ~ São José do~ SP
##  3       143          1 2010-09-15 0~ Av. Brigadeiro ~ São José do~ SP
##  4       195          1 2011-05-06 0~ Av. Brigadeiro ~ São José do~ SP
##  5       316          1 2012-10-27 0~ Av. Brigadeiro ~ São José do~ SP
##  6       327          1 2012-12-07 0~ Av. Brigadeiro ~ São José do~ SP
##  7       382          1 2013-08-07 0~ Av. Brigadeiro ~ São José do~ SP
##  8         1          2 2009-01-01 0~ Theodor-Heuss-S~ Stuttgart    <NA>
##  9        12          2 2009-02-11 0~ Theodor-Heuss-S~ Stuttgart    <NA>
## 10        67          2 2009-10-12 0~ Theodor-Heuss-S~ Stuttgart    <NA>
## # ... with more rows, and 15 more variables: BillingCountry <chr>,
## #   BillingPostalCode <chr>, Total <dbl>, FirstName <chr>, LastName <chr>,
## #   Company <chr>, Address <chr>, City <chr>, State <chr>, Country <chr>,
## #   PostalCode <chr>, Phone <chr>, Fax <chr>, Email <chr>, SupportRepId <int>
```

Part 4. Querying an API

1. Use `GET` from the `httr` package to return information about a topic that you are interested in that has been in the news lately. Select only the source, author, and title columns and print the tibble out.

```
#install.packages("httr")
library("httr")
```

```
## Warning: package 'httr' was built under R version 4.1.2
```

```
library(jsonlite)
```

`GET` is a function that does it what it says. It GETs a url. It is from the `httr` package. The URL we use is from the https://newsapi.org. Every API call is going to be unique. For newapi, there are two major endpoints : Everything and Top Headlines. We will use Top Headlines and specify the following details in the URL :

- country : us
- category : business, health, entertainment
- from : 2022-09-01
- language : en
- API key : which is generated uniquely for each user

```
business <- GET("https://newsapi.org/v2/top-headlines/?country=us&category=business&from=2022-09-01&lang
business_parsed <- fromJSON(rawToChar(business$content))
str(business_parsed, max.level = 1)
```

```
## List of 3
##  $ status      : chr "ok"
##  $ totalResults: int 59
##  $ articles    :'data.frame':    20 obs. of  8 variables:
```

```
business_parsed$articles %>%
  select(source, author, title)
```

```
##                   source.id             source.name           author
## 1                       <NA>                    CNBC       Abigail Ng
## 2                       <NA>                    CNBC       Jim Cramer
## 3                    reuters                 Reuters             <NA>
## 4                       <NA>              Daily Mail   Ruth Bashinsky
## 5           financial-times         Financial Times   Nicholas Megaw
## 6                       <NA>              MarketWatch    William Watts
## 7                       <NA>            Cointelegraph  Rakesh Upadhyay
## 8                       <NA>                    CNBC          Reuters
## 9                       <NA>                    CNBC  Brett Holzhauer
## 10                      <NA>                  SFGate       Katie Dowd
## 11                       cnn                     CNN      Zoe Sottile
## 12                      <NA>                 Decrypt        Tim Hakki
## 13                      <NA>              MarketWatch      Isabel Wang
## 14                      <NA>              Daily Mail      Alex Hammer
## 15                      <NA>            New York Post     Steve Cuozzo
## 16                      <NA>            New York Post Lisa Fickenscher
## 17        business-insider          Business Insider     Hannah Towey
## 18 the-wall-street-journal The Wall Street Journal     Aaron Tilley
## 19        business-insider          Business Insider    Sam Tabahriti
## 20                      <NA>            Cointelegraph       Yashu Gola
##
## 1                              Australia set to open flat; Fed, Bank of Japan rate decisions ahead t
## 2                           Jim Cramer: My biggest worry with the Fed and why it has us in a holding
## 3                                                 Frugal is the new cool for young Chinese as economy fa
## 4   Pennsylvania restaurant sue's customer who left $3,000 tip for waitress, but then failed to pay
## 5                              Market downturn sparks longest US tech IPO drought in over 20 years -
## 6                           Stock market's June lows are back in sight after S&P 500 loses grip on 3,90
## 7                          Here is why a 0.75% Fed rate hike could be bullish for Bitcoin and altcoins
## 8                            Volkswagen targets $70.1 billion to $75.1 billion valuation in planned Por
## 9                                  Should I buy stocks now or wait? Two experts weigh in on the curren
## 10                                              'Very limited' transbay BART service as 2 trains brea
## 11                                                    Google mistakenly sent an engineer almost
## 12  This Week on Crypto Twitter: Ethereum Merges, Hoskinson Gets Salty, Concerns Over Centralized Sta
## 13 Can the Fed tame inflation without further crushing the stock market? What investors need to know
```

```
## 14       Honda blasted for ordering hundreds of workers at Ohio factory to REPAY part of their bonus
## 15                 Back-to-work barometer falls short of measuring up to reality as offices fill up -
## 16                 Lawsuit that may have played role in Bed Bath & Beyond exec's suicide hits snags -
## 17                                    Travelers slam lengthy Airbnb chore lists and cleaning fees - Bu
## 18                                 Why Adobe Wants Figma and Why Some Investors Are Worried - The Wall
## 19      A man borrowed $75,000 for leg-lengthening surgery to make him 3 inches taller, report says
## 20                     Goldman Sachs' bearish macro outlook puts Bitcoin at risk of crashing to $12K
```

```
health <- GET("https://newsapi.org/v2/top-headlines/?country=us&category=health&from=2022-09-01&languag
health_parsed <- fromJSON(rawToChar(health$content))
str(health_parsed, max.level = 1)
```

```
## List of 3
##  $ status      : chr "ok"
##  $ totalResults: int 20
##  $ articles    :'data.frame':    20 obs. of  8 variables:
```

```
health_parsed$articles %>%
  select(source, author, title)
```

```
##              source.id          source.name                  author
## 1                 <NA>          Daily Beast               David Axe
## 2                 <NA>              KSL.com Emily Ashcraft, KSL.com
## 3                 <NA>             Medscape            Sara Freeman
## 4                 <NA>              INSIDER    Pocharapon Neammanee
## 5                 <NA> Neurosciencenews.com       Neuroscience News
## 6                 <NA>           Daily Mail     Mansur Shaheen U.S.
## 7             fox-news             Fox News           Shiv Sudhakar
## 8                 <NA> Neurosciencenews.com       Neuroscience News
## 9  the-washington-post  The Washington Post            Lisa Mulcahy
## 10                <NA>         SciTechDaily                    <NA>
## 11                <NA>              Express         Diana Buntajova
## 12                <NA>            Best Life              Adam Meyer
## 13                <NA>              nejm.org                    <NA>
## 14                <NA>          Queerty.com            David Hudson
## 15                <NA>              YouTube                    <NA>
## 16                <NA>          WSB Atlanta         Kirstin Garriss
## 17                <NA>          Eatthis.com               Desirée O
## 18                <NA>         Live Science        Nicoletta Lanese
## 19                <NA>         ScienceAlert             David Nield
## 20                 cnn                  CNN          Kristen Rogers
##
## 1                       Scientists Warn of Spike in Long COVID Cases Across the United States - The
## 2               Utah mom reflects on radioactive iodine cancer treatment, which has stood test of ti
## 3                                  Experts Express Caution Over Type 2 Diabetes-Tea Drinking Clai
## 4                     Family is fundraising for lawyer after they say police killed their pet racco
## 5                                  Adults Show Poorer Cognition, Better Well-Being with Age - Neuros
## 6                Third week of September is worst week of the year for asthma and allergy sufferers
## 7        Severe common cold cases increasing among young children may be pegged to COVID-19 lockdowns
## 8           Seven Healthy Lifestyle Habits May Reduce Dementia Risk for People With Diabetes - Neuros
## 9                                       What you need to know before you try a headstand - The Wash
## 10                                             5 Terrible Eating Habits That Cause Weight Gain -
## 11                          Pancreatic cancer: The 'gnawing' sign that can signal a tumour spreadi
```

```
## 12                                  Snoring Makes Your Cancer Risk Soar, Research Finds - Best Life
## 13                  Cerebral Embolic Protection during Transcatheter Aortic-Valve Replacement | NEJ
## 14          Gay men are the "canary in the coal mine" of future pandemics, warns doc seeking HIV cu
## 15                                 Harvard Nutritionist: The #1 Vitamin To Keep Your Brain Sharp - G
## 16    Doctors say some people may be 'COVID Super-dodgers' since they haven't gotten the virus yet -
## 17         The Worst Eating Habit for Colon Cancer, New Study Suggests - Eat This Not That - Eat Thi
## 18 In a 1st, scientists use designer immune cells to send an autoimmune disease into remission - Liv
## 19             There's One Simple Strategy to Reduce Alcohol Intake, Scientists Say, And It Works - S
## 20                                                          The 4-7-8 method that could help you
```

```
technology <- GET("https://newsapi.org/v2/top-headlines/?country=us&category=technology&from=2022-09-01
technology_parsed <- fromJSON(rawToChar(technology$content))
str(technology_parsed, max.level = 1)
```

```
## List of 3
##  $ status      : chr "ok"
##  $ totalResults: int 65
##  $ articles    :'data.frame':    20 obs. of  8 variables:
```

```
technology_parsed$articles %>%
  select(source, author, title)
```

```
##    source.id        source.name
## 1       <NA>            Gematsu
## 2       <NA>            Gematsu
## 3       <NA>     9to5google.com
## 4       <NA>             Kotaku
## 5   engadget           Engadget
## 6        ign                IGN
## 7       <NA>             Forbes
## 8        ign                IGN
## 9       <NA> Anime News Network
## 10      <NA>           Wccftech
## 11      <NA>   RoadandTrack.com
## 12       ign                IGN
## 13      <NA>           Wccftech
## 14      <NA>      Eurogamer.net
## 15      <NA>            9to5Mac
## 16      <NA>            Gematsu
## 17      <NA>      Nintendo Life
## 18      <NA>      Eurogamer.net
## 19      <NA>          MacRumors
## 20      <NA>            9to5Mac
##                                                 author
## 1                                         Sal Romano
## 2                                         Sal Romano
## 3                                           Abner Li
## 4                                       Luke Plunkett
## 5  https://www.engadget.com/about/editors/igor-bonifacic
## 6                                         Logan Plant
## 7                                           Erik Kain
## 8                                       IGN Japan Staff
## 9                                                 <NA>
```

```
## 10                                    Jason R. Wilson
## 11                                         Fred Smith
## 12                                    IGN Japan Staff
## 13                                   Francesco De Meo
## 14                                        Vikki Blake
## 15                                      Chance Miller
## 16                                         Sal Romano
## 17                                     Ollie Reynolds
## 18                                        Vikki Blake
## 19                                         Sami Fathi
## 20                                         José Adorno
##
## 1                                  Next Final Fantasy XVI trailer likely due out in October, says pro
## 2  Infinity Strash: Dragon Quest The Adventure of Dai details game modes, growth system; TGS 2022 gam
## 3                          Google ‘Nest Wifi Pro’ leaks with Wi-Fi 6E, higher price, and no point un
## 4                         This Massive Possible ’Grand Theft Auto VI’ Leak Sure Is Looking Legit [UP
## 5                            All iPhone 15 models will reportedly feature Dynamic Island display cut
## 6                                                    Tokyo Game Show 2022: Everything Announ
## 7  ‘Fortnite’ Season 4 Is Live - Here’s Everything New Including Map Changes, Battle Pass, Trailer An
## 8                    Resident Evil Village’s Shadows of Rose DLC Will ’Conclude the Winters Family Sa
## 9                                                  Tokyo Game Show 2022 - Photo Gallery - Ani
## 10     AMD Brings Ray Tracing Support In Its Open-Source Vulkan Graphics Drivers For RDNA 2 GPUs In M
## 11                                                 Listen to the 2024 Mustang Dark Horse Accelerate
## 12                                               La Mulana Director Wins Konami Contest to Revive Old
## 13                                              Diablo IV Gets Almost One Hour of Leaked Gameplay Foc
## 14                Sega has formally dropped the Yakuza brand name and replaced it with Like a Dragon
## 15                           iPhone 14 Pro camera shaking and rattling in TikTok, Snapchat, and other
## 16                                                                 Honkai: Star Rail ‘Witness’ tr
## 17                                                  Best Contra Games On Nintendo Systems
## 18           MMO Wizard101 is taken offline after an unhappy developer filled it with angry messages
## 19                                  Apple Investigating iPhone 14 Pro Models Freezing After Data Transf
## 20                        Apple October Event: New iPad Pro, iPad 10, M2 Macs, iPadOS 16, and macOS Ve
```