Activity 3

Dianah Marie Canonicato

2024-03-16

```r
#install.packages("dplyr")
#install.packages("rvest")
#install.packages("polite")
#install.packages("httr")

library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 4.3.3

##
## 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
```

```r
library(rvest)
```

```
## Warning: package 'rvest' was built under R version 4.3.3
```

```r
library(polite)
```

```
## Warning: package 'polite' was built under R version 4.3.3
```

```r
library(httr)
```

```
## Warning: package 'httr' was built under R version 4.3.3
```

```r
reviews <- data.frame()

link <- "https://www.imdb.com/title/tt0903747/reviews?ref_=tt_urv"
session <- bow(link, user_agent = "Educational Purpose")

scrapeNodes <- function(selector) {
```

```
  scrape(session) %>%
    html_nodes(selector) %>%
    html_text(trim = TRUE)
}

titleMovie <- scrapeNodes("a.subnav_heading")
reviewer <- scrapeNodes("span.display-name-link")[1:10]
scrapedDate <- scrapeNodes("span.review-date")[1:10]
scrapedRating <- scrapeNodes("span.rating-other-user-rating")[1:10]
scrapedReview <- scrapeNodes("div.text.show-more__control")[1:10]
scrapedTitleReview <- scrapeNodes("a.title")[1:10]

reviews <- rbind(reviews, data.frame(
  movie_title = rep(titleMovie, 10),
  reviewer_name = reviewer,
  reviewer_date = scrapedDate,
  review_rating = scrapedRating,
  review_content = scrapedReview,
  review_title = scrapedTitleReview
))

reviews
```

```
##       movie_title      reviewer_name    reviewer_date review_rating
## 1  Breaking Bad            FiRE010       4 July 2021          10/10
## 2  Breaking Bad    Supermanfan-13   9 November 2021          10/10
## 3  Breaking Bad       DiCaprioFan1   9 December 2022          10/10
## 4  Breaking Bad  TheLittleSongbird 13 November 2017          10/10
## 5  Breaking Bad       KinoKoopaKid      30 July 2021          10/10
## 6  Breaking Bad        gogoschka-1  12 January 2014          10/10
## 7  Breaking Bad       jehuschultz 19 February 2020          10/10
## 8  Breaking Bad         bruhperson      7 March 2019          10/10
## 9  Breaking Bad     Leofwine_draca       5 May 2021          10/10
## 10 Breaking Bad dhanushreddy-14919     18 July 2021          10/10
##
## 1
## 2
## 3
## 4  'Breaking Bad' is one of the most popular rated shows on IMDb, is one of those rarities where eve
## 5
## 6
## 7
## 8
## 9
## 10
##
## 1
## 2
## 3
## 4                                                          Among the best and most ac
## 5
## 6  If you mix Scarface, Robin Hood and maybe Tyler Durden with enough meth - you'll get a mean cockta
## 7
```

```
## 8
## 9                                                            Every bit a
## 10
```

// creating connections

```r
library(dplyr,dbplyr)
library(RMySQL)
```

```
## Warning: package 'RMySQL' was built under R version 4.3.3
```

```
## Loading required package: DBI
```

```
## Warning: package 'DBI' was built under R version 4.3.2
```

```r
#creating connections
connection <- dbConnect(RMySQL::MySQL(),
                        dsn="MariaDB-connection",
                        Server = "localhost",
                        dbname = "canonicato_2c",
                        user = "root",
                        password = "")
```

```r
dbListTables(connection)
```

```
## [1] "arvixtable" "lab2movies" "reviews"
```

```r
dbListFields(connection,"reviews")
```

```
## [1] "id"            "movie_title"    "reviewer_name"  "reviewer_date"
## [5] "review_rating"  "review_content" "review_title"
```

// reading data from database

```r
review_data <- dbGetQuery(connection, "SELECT * FROM canonicato_2c.reviews")
```

```r
glimpse(review_data)
```

```
## Rows: 50
## Columns: 7
## $ id            <int> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, ~
## $ movie_title    <chr> "Breaking Bad", "Breaking Bad", "Breaking Bad", "Breaki~
## $ reviewer_name  <chr> "FiRE010", "Supermanfan-13", "DiCaprioFan1", "TheLittle~
## $ reviewer_date  <chr> "4 July 2021", "9 November 2021", "9 December 2022", "1~
## $ review_rating  <chr> "10/10", "10/10", "10/10", "10/10", "10/10", "10/10", "~
## $ review_content <chr> "I have never watched a show that is as consistently ge~
## $ review_title   <chr> "Really Great", "Damn near perfect!", "Absolutely Brill~
```

// inserting reviews

```r
for (i in 1:nrow(reviews)) {
    query <- paste0("INSERT INTO reviews (movie_title, reviewer_name, reviewer_date, review_rating, rev
                VALUES ('", reviews$movie_title[i],"','", reviews$reviewer_name[i], "','", reviews$re
                gsub("'","''", reviews$review_content[i]), "','", gsub("'","''", reviews$review_title

  # execute query
  query_result <- dbSendQuery(connection, query)

  # clear result after executing query
  dbClearResult(query_result)
}
```

// read and view

```r
review_data <- dbGetQuery(connection, "SELECT * FROM canonicato_2c.reviews")


# VIEW REVIEWS without primary key "id"
review_data[,c(2:7)]
```

```
##       movie_title      reviewer_name      reviewer_date review_rating
## 1   Breaking Bad             FiRE010        4 July 2021         10/10
## 2   Breaking Bad       Supermanfan-13   9 November 2021         10/10
## 3   Breaking Bad         DiCaprioFan1   9 December 2022         10/10
## 4   Breaking Bad    TheLittleSongbird 13 November 2017         10/10
## 5   Breaking Bad          KinoKoopaKid       30 July 2021         10/10
## 6   Breaking Bad          jehuschultz 19 February 2020         10/10
## 7   Breaking Bad           gogoschka-1  12 January 2014         10/10
## 8   Breaking Bad            bruhperson       7 March 2019         10/10
## 9   Breaking Bad        Leofwine_draca        5 May 2021         10/10
## 10  Breaking Bad dhanushreddy-14919      18 July 2021         10/10
## 11  Breaking Bad             FiRE010        4 July 2021         10/10
## 12  Breaking Bad       Supermanfan-13   9 November 2021         10/10
## 13  Breaking Bad         DiCaprioFan1   9 December 2022         10/10
## 14  Breaking Bad    TheLittleSongbird 13 November 2017         10/10
## 15  Breaking Bad          KinoKoopaKid       30 July 2021         10/10
## 16  Breaking Bad          jehuschultz 19 February 2020         10/10
## 17  Breaking Bad           gogoschka-1  12 January 2014         10/10
## 18  Breaking Bad            bruhperson       7 March 2019         10/10
## 19  Breaking Bad        Leofwine_draca        5 May 2021         10/10
## 20  Breaking Bad dhanushreddy-14919      18 July 2021         10/10
## 21  Breaking Bad             FiRE010        4 July 2021         10/10
## 22  Breaking Bad       Supermanfan-13   9 November 2021         10/10
## 23  Breaking Bad         DiCaprioFan1   9 December 2022         10/10
## 24  Breaking Bad    TheLittleSongbird 13 November 2017         10/10
## 25  Breaking Bad          KinoKoopaKid       30 July 2021         10/10
## 26  Breaking Bad          jehuschultz 19 February 2020         10/10
## 27  Breaking Bad           gogoschka-1  12 January 2014         10/10
## 28  Breaking Bad            bruhperson       7 March 2019         10/10
## 29  Breaking Bad        Leofwine_draca        5 May 2021         10/10
## 30  Breaking Bad dhanushreddy-14919      18 July 2021         10/10
## 31  Breaking Bad             FiRE010        4 July 2021         10/10
## 32  Breaking Bad       Supermanfan-13   9 November 2021         10/10
```

```
## 33 Breaking Bad        DiCaprioFan1  9 December 2022         10/10
## 34 Breaking Bad  TheLittleSongbird 13 November 2017         10/10
## 35 Breaking Bad        KinoKoopaKid    30 July 2021         10/10
## 36 Breaking Bad         jehuschultz 19 February 2020         10/10
## 37 Breaking Bad         gogoschka-1  12 January 2014         10/10
## 38 Breaking Bad           bruhperson     7 March 2019         10/10
## 39 Breaking Bad      Leofwine_draca      5 May 2021         10/10
## 40 Breaking Bad dhanushreddy-14919     18 July 2021         10/10
## 41 Breaking Bad             FiRE010      4 July 2021         10/10
## 42 Breaking Bad      Supermanfan-13  9 November 2021         10/10
## 43 Breaking Bad        DiCaprioFan1  9 December 2022         10/10
## 44 Breaking Bad  TheLittleSongbird 13 November 2017         10/10
## 45 Breaking Bad        KinoKoopaKid    30 July 2021         10/10
## 46 Breaking Bad         jehuschultz 19 February 2020         10/10
## 47 Breaking Bad         gogoschka-1  12 January 2014         10/10
## 48 Breaking Bad           bruhperson     7 March 2019         10/10
## 49 Breaking Bad      Leofwine_draca      5 May 2021         10/10
## 50 Breaking Bad dhanushreddy-14919     18 July 2021         10/10
## 51 Breaking Bad             FiRE010      4 July 2021         10/10
## 52 Breaking Bad      Supermanfan-13  9 November 2021         10/10
## 53 Breaking Bad        DiCaprioFan1  9 December 2022         10/10
## 54 Breaking Bad  TheLittleSongbird 13 November 2017         10/10
## 55 Breaking Bad        KinoKoopaKid    30 July 2021         10/10
## 56 Breaking Bad         gogoschka-1  12 January 2014         10/10
## 57 Breaking Bad         jehuschultz 19 February 2020         10/10
## 58 Breaking Bad           bruhperson     7 March 2019         10/10
## 59 Breaking Bad      Leofwine_draca      5 May 2021         10/10
## 60 Breaking Bad dhanushreddy-14919     18 July 2021         10/10
##
## 1
## 2
## 3
## 4  'Breaking Bad' is one of the most popular rated shows on IMDb, is one of those rarities where eve
## 5
## 6
## 7
## 8
## 9
## 10
## 11
## 12
## 13
## 14 'Breaking Bad' is one of the most popular rated shows on IMDb, is one of those rarities where eve
## 15
## 16
## 17
## 18
## 19
## 20
## 21
## 22
## 23
## 24 'Breaking Bad' is one of the most popular rated shows on IMDb, is one of those rarities where eve
## 25
```

```
## 26
## 27
## 28
## 29
## 30
## 31
## 32
## 33
## 34 'Breaking Bad' is one of the most popular rated shows on IMDb, is one of those rarities where eve
## 35
## 36
## 37
## 38
## 39
## 40
## 41
## 42
## 43
## 44 'Breaking Bad' is one of the most popular rated shows on IMDb, is one of those rarities where eve
## 45
## 46
## 47
## 48
## 49
## 50
## 51
## 52
## 53
## 54 'Breaking Bad' is one of the most popular rated shows on IMDb, is one of those rarities where eve
## 55
## 56
## 57
## 58
## 59
## 60
## 
## 1
## 2
## 3
## 4                                                            Among the best and most ac
## 5
## 6
## 7  If you mix Scarface, Robin Hood and maybe Tyler Durden with enough meth - you'll get a mean cockta
## 8
## 9                                                                       Every bit a
## 10                                                                             A
## 11
## 12
## 13
## 14                                                           Among the best and most ac
## 15
## 16
## 17 If you mix Scarface, Robin Hood and maybe Tyler Durden with enough meth - you'll get a mean cockta
## 18
```

```
## 19                                                             Every bit a
## 20                                                                        A
## 21
## 22
## 23
## 24                                              Among the best and most ac
## 25
## 26
## 27 If you mix Scarface, Robin Hood and maybe Tyler Durden with enough meth - you'll get a mean cockta
## 28
## 29                                                             Every bit a
## 30                                                                        A
## 31
## 32
## 33
## 34                                              Among the best and most ac
## 35
## 36
## 37 If you mix Scarface, Robin Hood and maybe Tyler Durden with enough meth - you'll get a mean cockta
## 38
## 39                                                             Every bit a
## 40                                                                        A
## 41
## 42
## 43
## 44                                              Among the best and most ac
## 45
## 46
## 47 If you mix Scarface, Robin Hood and maybe Tyler Durden with enough meth - you'll get a mean cockta
## 48
## 49                                                             Every bit a
## 50                                                                        A
## 51
## 52
## 53
## 54                                              Among the best and most ac
## 55
## 56 If you mix Scarface, Robin Hood and maybe Tyler Durden with enough meth - you'll get a mean cockta
## 57
## 58
## 59                                                             Every bit a
## 60                                                                        A
```

```r
#View(review_data[,c(2:7)])

# Export to csv
write.csv(review_data[,c(2:7)], "LabEx3.csv")
```

// closing the database connections

```r
dbDisconnect(connection)
```

```
## [1] TRUE
```