CS102-IT2C Fermano Lab3.Rmd

Fermano, Aaron John

2024-03-14

```
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(rvest)
library(polite)
library(httr)
library(selectr)
```

```
movierevs = data.frame()
url <- "https://www.imdb.com/title/tt11564570/reviews?ref_=tt_urv"

 session <- bow(url,
                user_agent = "For Educational Purpose")

  scrapeNodes <- function(selector){
    scrape(session) %>%
      html_nodes(selector) %>%
      html_text(trim = TRUE)
  }
 movietitle <- rep("Glass Onion: A Knives Out Mystery",10)

  reviewer <- scrapeNodes("span.display-name-link")
  reviewer <- reviewer[1:10]

  ratingtitle <- scrapeNodes("a.title")
  ratingtitle <- ratingtitle[1:10]

  moviereviews <- scrapeNodes("div.content")
  moviereviews <- moviereviews[1:10]

  reviewdate <- scrapeNodes("span.review-date")
  reviewdate <- reviewdate[1:10]

  reviewratings <- scrapeNodes("div.ipl-ratings-bar")
```

```r
reviewratings <- reviewratings[1:10]

movierevs <- rbind(movierevs, data.frame(movie = movietitle,
                                         name = reviewer,
                                         title = ratingtitle,
                                         reviews = moviereviews,
                                         date = reviewdate,
                                         ratings = reviewratings))

write.csv(movierevs, file = "MovieGlassOnionRevs.csv")
```