CS102_JUNTANILLA_LAB3

JOSE ROLAND JUNTANILLA

2024-03-13

```
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)
```

Movie Reviews

```
reviewsdf = data.frame()

urllink <- "https://www.imdb.com/title/tt15239678/reviews?ref_=tt_urv"

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


  scrapemovie <- function(selector){
    scrape(session) %>%
      html_nodes(selector) %>%
      html_text(trim = TRUE)
  }

  movtitle <- rep("Dune: Part Two",10)

  revname <- scrapemovie("span.display-name-link")
  revname <- revname[1:10]

  ratetitle<- scrapemovie("a.title")
  ratetitle <- ratetitle[1:10]

  movreviews <- scrapemovie("div.content")
  movreviews <- movreviews[1:10]

  dateofrev <- scrapemovie("span.review-date")
```

```
dateofrev <- dateofrev[1:10]

revrate <- scrapemovie("div.ipl-ratings-bar")
revrate<- revrate[1:10]



Scrapereviews <- rbind(reviewsdf, data.frame(Movie = movtitle,
                    "Name of Reviewer" = revname,
                     Title = ratetitle,
                     Reviews = movreviews,
                    "Date of Reviews" = dateofrev,
                     Ratings = revrate))

write.csv(Scrapereviews, file = "Reviewsdatabase.csv")
```