

# lab\_exersice#1\_Echaveria

Luigi Echaveria

2024-02-12

## SCRAPING FOR PILLOWS

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

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)

pillow_products <- data.frame()

pillowlink1 = "https://www.amazon.co.uk/s?k=pillows&crid=3CYPNR2J713S1&sprefix=pillows%2Caps%2C717&ref="

session1 <- bow(pillowlink1,
               user_agent = "Educational Purpose")

scrapeNodes <- function(selector){
  scrape(session1) %>%
    html_nodes(selector) %>%
    html_text(trim = TRUE)
}

scrapedCategory <- "Pillow"

scrapedName <- scrapeNodes("h2.a-size-mini.a-spacing-none.a-color-base.s-line-clamp-4")
scrapedName <- scrapedName[1:45]

scrapedPrice <- scrapeNodes("span.a-offscreen")
scrapedPrice <- scrapedPrice[1:45]
```

```

scrapedRatings <- scrapeNodes("span.a-icon-alt")
scrapedRatings <- scrapedRatings[1:45]

scrapedReviews <- scrapeNodes("span.a-size-base.s-underline-text")
scrapedReviews <- scrapedReviews[1:45]

pillow_products <- rbind(pillow_products, data.frame(category = scrapedCategory, name = scrapedName,
  price = scrapedPrice,
  ratings = scrapedRatings,
  no_of_reviews = scrapedReviews))

#####

pillowlink2 = "https://www.amazon.co.uk/s?k=pillows&page=2&crd=3CYPNR2J713S1&qid=1710230571&srefix=pi

session2 <- bow(pillowlink2,
  user_agent = "For Educational Purpose")

scrapedCategory2 <- "Pillow"

scrapedName2 <- scrapeNodes("h2.a-size-mini.a-spacing-none.a-color-base.s-line-clamp-4")
scrapedName2 <- scrapedName2[1:45]

scrapedPrice2 <- scrapeNodes("span.a-offscreen")
scrapedPrice2 <- scrapedPrice2[1:45]

scrapedRatings2 <- scrapeNodes("span.a-icon-alt")
scrapedRatings2 <- scrapedRatings2[1:45]

scrapedReviews2 <- scrapeNodes("span.a-size-base.s-underline-text")
scrapedReviews2 <- scrapedReviews2[1:45]

pillow_products <- rbind(pillow_products, data.frame(category = scrapedCategory, name = scrapedName,
  price = scrapedPrice,
  ratings = scrapedRatings,
  no_of_reviews = scrapedReviews))

#####

pillowlink3 = "https://www.amazon.co.uk/s?k=pillows&page=3&crd=3CYPNR2J713S1&qid=1710233098&srefix=pi

session3 <- bow(pillowlink3,
  user_agent = "For Educational Purpose")

scrapedCategory3 <- "Pillow"

scrapedName3 <- scrapeNodes("h2.a-size-mini.a-spacing-none.a-color-base.s-line-clamp-4")
scrapedName3 <- scrapedName3[1:45]

```

```

scrapedPrice3 <- scrapeNodes("span.a-offscreen")
scrapedPrice3 <- scrapedPrice3[1:45]

scrapedRatings3 <- scrapeNodes("span.a-icon-alt")
scrapedRatings3 <- scrapedRatings3[1:45]

scrapedReviews3 <- scrapeNodes("span.a-size-base.s-underline-text")
scrapedReviews3 <- scrapedReviews3[1:45]

pillow_products <- rbind(pillow_products, data.frame(category = scrapedCategory, name = scrapedName,
  price = scrapedPrice,
  ratings = scrapedRatings,
  no_of_reviews = scrapedReviews))

pillow_products <- pillow_products[1:100,]

#write.csv(pillow_products, "pillow_products.csv")
#View(pillow_products)

```

## SCRAPING FOR SOCKS

```

library(dplyr)
library(rvest)
library(polite)
library(httr)

socks_products <- data.frame()

sockslink1_1_1 = "https://www.amazon.co.uk/s?k=socks&crd=8MCF9HUOG7H0&qid=1710235362&srefix=socks%2Cap

session4_4_4 <- bow(sockslink1_1_1,
  user_agent = "Education")

scrapeNodes <- function(selector){
  scrape(session4_4_4) %>%
    html_nodes(selector) %>%
    html_text(trim = TRUE)
}

scrapedCategory4 <- "Socks"

scrapedsocksName <- scrapeNodes("h2.a-size-mini.a-spacing-none.a-color-base.s-line-clamp-4")
scrapedsocksName <- scrapedsocksName[1:45]

scrapedsocksPrice <- scrapeNodes("span.a-offscreen")
scrapedsocksPrice <- scrapedsocksPrice[1:45]

scrapedsocksRatings <- scrapeNodes("span.a-icon-alt")
scrapedsocksRatings <- scrapedsocksRatings[1:45]

scrapedsocksReviews <- scrapeNodes("span.a-size-base.s-underline-text")
scrapedsocksReviews <- scrapedsocksReviews[1:45]

```

```

socks_products <- rbind(socks_products, data.frame(category = scrapedCategory4, name = scrapedsocksName,
  price = scrapedsocksPrice,
  ratings = scrapedsocksRatings,
  no_of_reviews = scrapedsocksReviews))

#####

sockslink2_2_2 = "https://www.amazon.co.uk/s?k=socks&page=2&crid=8MCF9HU0G7H0&qid=1710235371&srefix=so

session5_5_5 <- bow(sockslink2_2_2,
  user_agent = "Educational Purpose")

scrapeNodes <- function(selector){
  scrape(session5_5_5) %>%
    html_nodes(selector) %>%
    html_text(trim = TRUE)
}

scrapedCategory5 <- "Socks"

scrapedsocksName2 <- scrapeNodes("h2.a-size-mini.a-spacing-none.a-color-base.s-line-clamp-4")
scrapedsocksName2 <- scrapedsocksName2[1:45]

scrapedsocksPrice2 <- scrapeNodes("span.a-offscreen")
scrapedsocksPrice2 <- scrapedsocksPrice2[1:45]

scrapedsocksRatings2 <- scrapeNodes("span.a-icon-alt")
scrapedsocksRatings2 <- scrapedsocksRatings2[1:45]

scrapedsocksReviews2 <- scrapeNodes("span.a-size-base.s-underline-text")
scrapedsocksReviews2 <- scrapedsocksReviews2[1:45]

socks_products <- rbind(socks_products, data.frame(category = scrapedCategory5, name = scrapedsocksName2,
  price = scrapedsocksPrice2,
  ratings = scrapedsocksRatings2,
  no_of_reviews = scrapedsocksReviews2))

#####

sockslink3_3_3 = "https://www.amazon.co.uk/s?k=socks&page=3&crid=8MCF9HU0G7H0&qid=1710235444&srefix=so

session6_6_6 <- bow(sockslink3_3_3,
  user_agent = "Educational Purpose")

scrapeNodes <- function(selector){
  scrape(session6_6_6) %>%

```

```

    html_nodes(selector) %>%
    html_text(trim = TRUE)
}

scrapedCategory6 <- "Socks"

scrapedsocksName3 <- scrapeNodes("h2.a-size-mini.a-spacing-none.a-color-base.s-line-clamp-4")
scrapedsocksName3 <- scrapedsocksName3[1:45]

scrapedsocksPrice3 <- scrapeNodes("span.a-offscreen")
scrapedsocksPrice3 <- scrapedsocksPrice3[1:45]

scrapedsocksRatings3 <- scrapeNodes("span.a-icon-alt")
scrapedsocksRatings3 <- scrapedsocksRatings3[1:45]

scrapedsocksReviews3 <- scrapeNodes("span.a-size-base.s-underline-text")
scrapedsocksReviews3 <- scrapedsocksReviews3[1:45]

socks_products <- rbind(socks_products, data.frame(category = scrapedCategory5, name = scrapedsocksName2,
    price = scrapedsocksPrice2,
    ratings = scrapedsocksRatings2,
    no_of_reviews = scrapedsocksReviews2))

socks_products <- socks_products[1:100,]
#write.csv(socks_products, "socks_products.csv")

#View(socks_products)

```

## MERGING TWO PRODUCTS

```

scraped_products <- rbind(pillow_products, socks_products)
#write.csv(scraped_products, "byCategory.csv")
#View(scraped_products)

```

## 10 PRODUCTS

product1

```

library(dplyr)
library(rvest)
library(polite)
library(httr)
library(selectr)

product1 = data.frame()
scrapedCategory <- "Pillow"

###

link = "https://www.amazon.co.uk/product-reviews/B01L97Z3V8/ref=cm_cr_arp_d_viewopt_sr?ie=UTF8&filter=

sessionProd1 <- bow(link,
    user_agent = "For Educational Purpose")

```

```

scrapeNodes <- function(selector){
  scrape(sessionProd1) %>%
    html_nodes(selector) %>%
    html_text(trim = TRUE)
}

scrapedprodName <- scrapeNodes("h1.a-size-large.a-text-ellipsis")

scrapedReviewer <- scrapeNodes("span.a-profile-name")
scrapedReviewer <- scrapedReviewer[1:10]

scrapedReview <- scrapeNodes("span.a-size-base.review-text.review-text-content")
scrapedReview <- scrapedReview[1:10]

scrapedDate <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
scrapedDate <- scrapedDate[1:10]

scrapedprodRating <- scrapeNodes("span.a-icon-alt")
scrapedprodRating <- scrapedprodRating[1:10]

product1 <- rbind(product1, data.frame(category = scrapedCategory, name = scrapedprodName,
  reviewer = scrapedReviewer,
  review = scrapedReview,
  date = scrapedDate,
  ratings = scrapedprodRating))

Sys.sleep(5)

#####

link2 = "https://www.amazon.co.uk/product-reviews/B01L97Z3V8/ref=cm_cr_getr_d_paging_btm_next_2?ie=UTF8"

sessionProd2 <- bow(link2,
  user_agent = "For Educational Purpose")

scrapedReviewer2 <- scrapeNodes("span.a-profile-name")
scrapedReviewer2 <- scrapedReviewer2[1:10]

scrapedReview2 <- scrapeNodes("span.a-size-base.review-text.review-text-content")
scrapedReview2 <- scrapedReview2[1:10]

scrapedDate2 <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
scrapedDate2 <- scrapedDate2[1:10]

```

```

scrapedprodRating2 <- scrapeNodes("span.a-icon-alt")
scrapedprodRating2 <- scrapedprodRating2[1:10]

product1 <- rbind(product1, data.frame(category = scrapedCategory ,name = scrapedprodName,
    reviewer = scrapedReviewer2,
    review = scrapedReview2,
    date = scrapedDate2,
    ratings = scrapedprodRating2))

Sys.sleep(5)

#####

link3 = "https://www.amazon.co.uk/product-reviews/B01L97Z3V8/ref=cm_cr_getr_d_paging_btm_next_3?i

sessionProd3 <- bow(link3,
    user_agent = "For Educational Purpose")

scrapedReviewer3 <- scrapeNodes("span.a-profile-name")
scrapedReviewer3 <- scrapedReviewer3[1:10]

scrapedReview3 <- scrapeNodes("span.a-size-base.review-text.review-text-content")
scrapedReview3 <- scrapedReview3[1:10]

scrapedDate3 <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
scrapedDate3 <- scrapedDate3[1:10]

scrapedprodRating3 <- scrapeNodes("span.a-icon-alt")
scrapedprodRating3 <- scrapedprodRating3[1:10]

product1 <- rbind(product1, data.frame(category = scrapedCategory ,name = scrapedprodName,
    reviewer = scrapedReviewer3,
    review = scrapedReview3,
    date = scrapedDate3,
    ratings = scrapedprodRating3))

Sys.sleep(5)

#View(product1)
#write.csv(product1, "product1.csv")

product2

library(dplyr)
library(rvest)
library(polite)
library(httr)
library(selectr)

product2 = data.frame()

```

```

scrapedCategory <- "Pillow"

#####

link = "https://www.amazon.co.uk/product-reviews/B01FMHF1EA/ref=cm_cr_arp_d_viewopt_sr?ie=UTF8&filter

sessionProd1 <- bow(link,
  user_agent = "For Educational Purpose")

scrapeNodes <- function(selector){
  scrape(sessionProd1) %>%
    html_nodes(selector) %>%
    html_text(trim = TRUE)
}

scrapedprodName <- scrapeNodes("h1.a-size-large.a-text-ellipsis")

scrapedReviewer <- scrapeNodes("span.a-profile-name")
scrapedReviewer <- scrapedReviewer[1:10]

scrapedReview <- scrapeNodes("span.a-size-base.review-text.review-text-content")
scrapedReview <- scrapedReview[1:10]

scrapedDate <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
scrapedDate <- scrapedDate[1:10]

scrapedprodRating <- scrapeNodes("span.a-icon-alt")
scrapedprodRating <- scrapedprodRating[1:10]

product2 <- rbind(product2, data.frame(category = scrapedCategory, name = scrapedprodName,
  reviewer = scrapedReviewer,
  review = scrapedReview,
  date = scrapedDate,
  ratings = scrapedprodRating))

Sys.sleep(5)
#####

link2 = "https://www.amazon.co.uk/product-reviews/B01FMHF1EA/ref=cm_cr_getr_d_paging_btm_next_2?i

sessionProd2 <- bow(link2,
  user_agent = "For Educational Purpose")

scrapedReviewer2 <- scrapeNodes("span.a-profile-name")
scrapedReviewer2 <- scrapedReviewer2[1:10]

scrapedReview2 <- scrapeNodes("span.a-size-base.review-text.review-text-content")

```



```

scrapedReview2 <- scrapedReview2[1:10]

scrapedDate2 <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
scrapedDate2 <- scrapedDate2[1:10]

scrapedprodRating2 <- scrapeNodes("span.a-icon-alt")
scrapedprodRating2 <- scrapedprodRating2[1:10]

product2 <- rbind(product2, data.frame(category = scrapedCategory ,name = scrapedprodName,
  reviewer = scrapedReviewer2,
  review = scrapedReview2,
  date = scrapedDate2,
  ratings = scrapedprodRating2))

Sys.sleep(5)
#####

link3 = "https://www.amazon.co.uk/product-reviews/B01FMHF1EA/ref=cm_cr_getr_d_paging_btm_next_3?i

sessionProd3 <- bow(link3,
  user_agent = "For Educational Purpose")

scrapedReviewer3 <- scrapeNodes("span.a-profile-name")
scrapedReviewer3 <- scrapedReviewer3[1:10]

scrapedReview3 <- scrapeNodes("span.a-size-base.review-text.review-text-content")
scrapedReview3 <- scrapedReview3[1:10]

scrapedDate3 <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
scrapedDate3 <- scrapedDate3[1:10]

scrapedprodRating3 <- scrapeNodes("span.a-icon-alt")
scrapedprodRating3 <- scrapedprodRating3[1:10]

product2 <- rbind(product2, data.frame(category = scrapedCategory ,name = scrapedprodName,
  reviewer = scrapedReviewer3,
  review = scrapedReview3,
  date = scrapedDate3,
  ratings = scrapedprodRating3))

Sys.sleep(5)
#View(product2)
#write.csv(product2, "product2.csv")

```

product3

```

library(dplyr)
library(rvest)
library(polite)
library(httr)
library(selectr)

product3 = data.frame()
scrapedCategory <- "Pillow"

#####

link = "https://www.amazon.co.uk/product-reviews/B016RAAUEM/ref=cm_cr_arp_d_viewopt_sr?ie=UTF8&filter

sessionProd1 <- bow(link,
  user_agent = "For Educational Purpose")

scrapeNodes <- function(selector){
  scrape(sessionProd1) %>%
    html_nodes(selector) %>%
    html_text(trim = TRUE)
}

scrapedprodName <- scrapeNodes("h1.a-size-large.a-text-ellipsis")

scrapedReviewer <- scrapeNodes("span.a-profile-name")
scrapedReviewer <- scrapedReviewer[1:10]

scrapedReview <- scrapeNodes("span.a-size-base.review-text.review-text-content")
scrapedReview <- scrapedReview[1:10]

scrapedDate <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
scrapedDate <- scrapedDate[1:10]

scrapedprodRating <- scrapeNodes("span.a-icon-alt")
scrapedprodRating <- scrapedprodRating[1:10]

product3 <- rbind(product3, data.frame(category = scrapedCategory, name = scrapedprodName,
  reviewer = scrapedReviewer,
  review = scrapedReview,
  date = scrapedDate,
  ratings = scrapedprodRating))

Sys.sleep(5)
#####

link2 = "https://www.amazon.co.uk/product-reviews/B016RAAUEM/ref=cm_cr_getr_d_paging_btm_next_2?i

sessionProd2 <- bow(link2,

```

```

        user_agent = "For Educational Purpose")

scrapedReviewer2 <- scrapeNodes("span.a-profile-name")
scrapedReviewer2 <- scrapedReviewer2[1:10]

scrapedReview2 <- scrapeNodes("span.a-size-base.review-text.review-text-content")
scrapedReview2 <- scrapedReview2[1:10]

scrapedDate2 <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
scrapedDate2 <- scrapedDate2[1:10]

scrapedprodRating2 <- scrapeNodes("span.a-icon-alt")
scrapedprodRating2 <- scrapedprodRating2[1:10]

product3 <- rbind(product3, data.frame(category = scrapedCategory ,name = scrapedprodName,
        reviewer = scrapedReviewer2,
        review = scrapedReview2,
        date = scrapedDate2,
        ratings = scrapedprodRating2))

Sys.sleep(5)
#####

link3 = "https://www.amazon.co.uk/product-reviews/B016RAAUEM/ref=cm_cr_getr_d_paging_btm_next_3?i

sessionProd3 <- bow(link3,
        user_agent = "For Educational Purpose")

scrapedReviewer3 <- scrapeNodes("span.a-profile-name")
scrapedReviewer3 <- scrapedReviewer3[1:10]

scrapedReview3 <- scrapeNodes("span.a-size-base.review-text.review-text-content")
scrapedReview3 <- scrapedReview3[1:10]

scrapedDate3 <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
scrapedDate3 <- scrapedDate3[1:10]

scrapedprodRating3 <- scrapeNodes("span.a-icon-alt")
scrapedprodRating3 <- scrapedprodRating3[1:10]

product3 <- rbind(product3, data.frame(category = scrapedCategory ,name = scrapedprodName,
        reviewer = scrapedReviewer3,
        review = scrapedReview3,
        date = scrapedDate3,

```

```

        ratings = scrapedprodRating3))

    Sys.sleep(5)
    #View(product3)
    #write.csv(product3, "product3.csv")

product4
library(dplyr)
library(rvest)
library(polite)
library(httr)
library(selectr)

product4 = data.frame()
scrapedCategory <- "Pillow"

#####

link = "https://www.amazon.co.uk/product-reviews/B01CR9IRWY/ref=cm_cr_arp_d_viewopt_sr?ie=UTF8&filter

    sessionProd1 <- bow(link,
        user_agent = "For Educational Purpose")

    scrapeNodes <- function(selector){
    scrape(sessionProd1) %>%
        html_nodes(selector) %>%
        html_text(trim = TRUE)
    }

    scrapedprodName <- scrapeNodes("h1.a-size-large.a-text-ellipsis")

    scrapedReviewer <- scrapeNodes("span.a-profile-name")
    scrapedReviewer <- scrapedReviewer[1:10]

    scrapedReview <- scrapeNodes("span.a-size-base.review-text.review-text-content")
    scrapedReview <- scrapedReview[1:10]

    scrapedDate <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
    scrapedDate <- scrapedDate[1:10]

    scrapedprodRating <- scrapeNodes("span.a-icon-alt")
    scrapedprodRating <- scrapedprodRating[1:10]

    product4 <- rbind(product4, data.frame(category = scrapedCategory, name = scrapedprodName,
        reviewer = scrapedReviewer,
        review = scrapedReview,
        date = scrapedDate,

```

```

        ratings = scrapedprodRating))

Sys.sleep(5)
#####

link2 = "https://www.amazon.co.uk/product-reviews/B01CR9IRWY/ref=cm_cr_getr_d_paging_btm_next_2?i

sessionProd2 <- bow(link2,
                    user_agent = "For Educational Purpose")

scrapedReviewer2 <- scrapeNodes("span.a-profile-name")
scrapedReviewer2 <- scrapedReviewer2[1:10]

scrapedReview2 <- scrapeNodes("span.a-size-base.review-text.review-text-content")
scrapedReview2 <- scrapedReview2[1:10]

scrapedDate2 <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
scrapedDate2 <- scrapedDate2[1:10]

scrapedprodRating2 <- scrapeNodes("span.a-icon-alt")
scrapedprodRating2 <- scrapedprodRating2[1:10]

product4 <- rbind(product4, data.frame(category = scrapedCategory ,name = scrapedprodName,
                                       reviewer = scrapedReviewer2,
                                       review = scrapedReview2,
                                       date = scrapedDate2,
                                       ratings = scrapedprodRating2))

Sys.sleep(5)
#####

link3 = "https://www.amazon.co.uk/product-reviews/B01CR9IRWY/ref=cm_cr_getr_d_paging_btm_next_3?i

sessionProd3 <- bow(link3,
                    user_agent = "For Educational Purpose")

scrapedReviewer3 <- scrapeNodes("span.a-profile-name")
scrapedReviewer3 <- scrapedReviewer3[1:10]

scrapedReview3 <- scrapeNodes("span.a-size-base.review-text.review-text-content")
scrapedReview3 <- scrapedReview3[1:10]

scrapedDate3 <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
scrapedDate3 <- scrapedDate3[1:10]

```

```

scrapedprodRating3 <- scrapeNodes("span.a-icon-alt")
scrapedprodRating3 <- scrapedprodRating3[1:10]

product4 <- rbind(product4, data.frame(category = scrapedCategory ,name = scrapedprodName,
                                     reviewer = scrapedReviewer3,
                                     review = scrapedReview3,
                                     date = scrapedDate3,
                                     ratings = scrapedprodRating3))

Sys.sleep(5)
#View(product4)
#write.csv(product4, "product4.csv")

product5
library(dplyr)
library(rvest)
library(polite)
library(httr)
library(selectr)

product5 = data.frame()
scrapedCategory <- "Pillows"

#####

link = "https://www.amazon.co.uk/product-reviews/B01G51NIJ2/ref=cm_cr_arp_d_viewopt_sr?ie=UTF8&filter

sessionProd1 <- bow(link,
                    user_agent = "For Educational Purpose")

scrapeNodes <- function(selector){
  scrape(sessionProd1) %>%
    html_nodes(selector) %>%
    html_text(trim = TRUE)
}

scrapedprodName <- scrapeNodes("h1.a-size-large.a-text-ellipsis")

scrapedReviewer <- scrapeNodes("span.a-profile-name")
scrapedReviewer <- scrapedReviewer[1:10]

scrapedReview <- scrapeNodes("span.a-size-base.review-text.review-text-content")
scrapedReview <- scrapedReview[1:10]

scrapedDate <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
scrapedDate <- scrapedDate[1:10]

```

```

scrapedprodRating <- scrapeNodes("span.a-icon-alt")
scrapedprodRating <- scrapedprodRating[1:10]

product5 <- rbind(product5, data.frame(category = scrapedCategory, name = scrapedprodName,
  reviewer = scrapedReviewer,
  review = scrapedReview,
  date = scrapedDate,
  ratings = scrapedprodRating))

Sys.sleep(5)
#####

link2 = "https://www.amazon.co.uk/product-reviews/B01G51NIJ2/ref=cm_cr_getr_d_paging_btm_next_2?i

sessionProd2 <- bow(link2,
  user_agent = "For Educational Purpose")

scrapedReviewer2 <- scrapeNodes("span.a-profile-name")
scrapedReviewer2 <- scrapedReviewer2[1:10]

scrapedReview2 <- scrapeNodes("span.a-size-base.review-text.review-text-content")
scrapedReview2 <- scrapedReview2[1:10]

scrapedDate2 <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
scrapedDate2 <- scrapedDate2[1:10]

scrapedprodRating2 <- scrapeNodes("span.a-icon-alt")
scrapedprodRating2 <- scrapedprodRating2[1:10]

product5 <- rbind(product5, data.frame(category = scrapedCategory ,name = scrapedprodName,
  reviewer = scrapedReviewer2,
  review = scrapedReview2,
  date = scrapedDate2,
  ratings = scrapedprodRating2))

Sys.sleep(5)
#####

link3 = "https://www.amazon.co.uk/product-reviews/B01G51NIJ2/ref=cm_cr_getr_d_paging_btm_next_3?i

sessionProd3 <- bow(link3,
  user_agent = "For Educational Purpose")

scrapedReviewer3 <- scrapeNodes("span.a-profile-name")
scrapedReviewer3 <- scrapedReviewer3[1:10]

```

```

scrapedReview3 <- scrapeNodes("span.a-size-base.review-text.review-text-content")
scrapedReview3 <- scrapedReview3[1:10]

scrapedDate3 <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
scrapedDate3 <- scrapedDate3[1:10]

scrapedprodRating3 <- scrapeNodes("span.a-icon-alt")
scrapedprodRating3 <- scrapedprodRating3[1:10]

product5 <- rbind(product5, data.frame(category = scrapedCategory ,name = scrapedprodName,
                                     reviewer = scrapedReviewer3,
                                     review = scrapedReview3,
                                     date = scrapedDate3,
                                     ratings = scrapedprodRating3))

Sys.sleep(5)

#View(product5)
#write.csv(product5, "product5.csv")

product6
library(dplyr)
library(rvest)
library(polite)
library(httr)
library(selectr)

product6 = data.frame()
scrapedCategory <- "Socks"

#####

link = "https://www.amazon.co.uk/product-reviews/B07FCCT6G6/ref=cm_cr_arp_d_viewopt_sr?ie=UTF8&filter

sessionProd1 <- bow(link,
                    user_agent = "For Educational Purpose")

scrapeNodes <- function(selector){
  scrape(sessionProd1) %>%
    html_nodes(selector) %>%
    html_text(trim = TRUE)
}

scrapedprodName <- scrapeNodes("h1.a-size-large.a-text-ellipsis")

scrapedReviewer <- scrapeNodes("span.a-profile-name")
scrapedReviewer <- scrapedReviewer[1:10]

```



```

scrapedReview <- scrapeNodes("span.a-size-base.review-text.review-text-content")
scrapedReview <- scrapedReview[1:10]

scrapedDate <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
scrapedDate <- scrapedDate[1:10]

scrapedprodRating <- scrapeNodes("span.a-icon-alt")
scrapedprodRating <- scrapedprodRating[1:10]

product6 <- rbind(product6, data.frame(category = scrapedCategory, name = scrapedprodName,
  reviewer = scrapedReviewer,
  review = scrapedReview,
  date = scrapedDate,
  ratings = scrapedprodRating))

#####

link2 = "https://www.amazon.co.uk/product-reviews/B07FCCT6G6/ref=cm_cr_getr_d_paging_btm_next_2?i

sessionProd2 <- bow(link2,
  user_agent = "For Educational Purpose")

scrapedReviewer2 <- scrapeNodes("span.a-profile-name")
scrapedReviewer2 <- scrapedReviewer2[1:10]

scrapedReview2 <- scrapeNodes("span.a-size-base.review-text.review-text-content")
scrapedReview2 <- scrapedReview2[1:10]

scrapedDate2 <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
scrapedDate2 <- scrapedDate2[1:10]

scrapedprodRating2 <- scrapeNodes("span.a-icon-alt")
scrapedprodRating2 <- scrapedprodRating2[1:10]

product6 <- rbind(product6, data.frame(category = scrapedCategory ,name = scrapedprodName,
  reviewer = scrapedReviewer2,
  review = scrapedReview2,
  date = scrapedDate2,
  ratings = scrapedprodRating2))

#####

link3 = "https://www.amazon.co.uk/product-reviews/B07FCCT6G6/ref=cm_cr_getr_d_paging_btm_next_3?i

sessionProd3 <- bow(link3,
  user_agent = "For Educational Purpose")

```

```

scrapedReviewer3 <- scrapeNodes("span.a-profile-name")
scrapedReviewer3 <- scrapedReviewer3[1:10]

scrapedReview3 <- scrapeNodes("span.a-size-base.review-text.review-text-content")
scrapedReview3 <- scrapedReview3[1:10]

scrapedDate3 <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
scrapedDate3 <- scrapedDate3[1:10]

scrapedprodRating3 <- scrapeNodes("span.a-icon-alt")
scrapedprodRating3 <- scrapedprodRating3[1:10]

product6 <- rbind(product6, data.frame(category = scrapedCategory ,name = scrapedprodName,
                                     reviewer = scrapedReviewer3,
                                     review = scrapedReview3,
                                     date = scrapedDate3,
                                     ratings = scrapedprodRating3))

Sys.sleep(5)

#View(product6)
#write.csv(product6, "product6.csv")

product7

library(dplyr)
library(rvest)
library(polite)
library(httr)
library(selectr)

product7 = data.frame()
scrapedCategory <- "Socks"

#####

link = "https://www.amazon.co.uk/product-reviews/B0150BE59U/ref=cm_cr_arp_d_viewopt_sr?ie=UTF8&filter

sessionProd1 <- bow(link,
                    user_agent = "For Educational Purpose")

scrapeNodes <- function(selector){
  scrape(sessionProd1) %>%
    html_nodes(selector) %>%
    html_text(trim = TRUE)
}

scrapedprodName <- scrapeNodes("h1.a-size-large.a-text-ellipsis")

```

```

scrapedReviewer <- scrapeNodes("span.a-profile-name")
scrapedReviewer <- scrapedReviewer[1:10]

scrapedReview <- scrapeNodes("span.a-size-base.review-text.review-text-content")
scrapedReview <- scrapedReview[1:10]

scrapedDate <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
scrapedDate <- scrapedDate[1:10]

scrapedprodRating <- scrapeNodes("span.a-icon-alt")
scrapedprodRating <- scrapedprodRating[1:10]

product7 <- rbind(product7, data.frame(category = scrapedCategory, name = scrapedprodName,
  reviewer = scrapedReviewer,
  review = scrapedReview,
  date = scrapedDate,
  ratings = scrapedprodRating))

Sys.sleep(5)
#####

link2 = "https://www.amazon.co.uk/product-reviews/B0150BE59U/ref=cm_cr_getr_d_paging_btm_next_2?i

sessionProd2 <- bow(link2,
  user_agent = "For Educational Purpose")

scrapedReviewer2 <- scrapeNodes("span.a-profile-name")
scrapedReviewer2 <- scrapedReviewer2[1:10]

scrapedReview2 <- scrapeNodes("span.a-size-base.review-text.review-text-content")
scrapedReview2 <- scrapedReview2[1:10]

scrapedDate2 <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
scrapedDate2 <- scrapedDate2[1:10]

scrapedprodRating2 <- scrapeNodes("span.a-icon-alt")
scrapedprodRating2 <- scrapedprodRating2[1:10]

product7 <- rbind(product7, data.frame(category = scrapedCategory ,name = scrapedprodName,
  reviewer = scrapedReviewer2,
  review = scrapedReview2,
  date = scrapedDate2,
  ratings = scrapedprodRating2))

Sys.sleep(5)
#####

```

```

link3 = "https://www.amazon.co.uk/product-reviews/B0150BE59U/ref=cm_cr_getr_d_paging_btm_next_3?ie=UTF8&filterBy=reviewer_rated&pageNumber=3"

sessionProd3 <- bow(link3,
  user_agent = "For Educational Purpose")

scrapedReviewer3 <- scrapeNodes("span.a-profile-name")
scrapedReviewer3 <- scrapedReviewer3[1:10]

scrapedReview3 <- scrapeNodes("span.a-size-base.review-text.review-text-content")
scrapedReview3 <- scrapedReview3[1:10]

scrapedDate3 <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
scrapedDate3 <- scrapedDate3[1:10]

scrapedprodRating3 <- scrapeNodes("span.a-icon-alt")
scrapedprodRating3 <- scrapedprodRating3[1:10]

product7 <- rbind(product7, data.frame(category = scrapedCategory ,name = scrapedprodName,
  reviewer = scrapedReviewer3,
  review = scrapedReview3,
  date = scrapedDate3,
  ratings = scrapedprodRating3))

Sys.sleep(5)

#View(product7)
#write.csv(product7, "product7.csv")

```

product8

```

library(dplyr)
library(rvest)
library(polite)
library(httr)
library(selectr)

product8 = data.frame()
scrapedCategory <- "Socks"

#####

link = "https://www.amazon.co.uk/product-reviews/B099FJCQSF/ref=cm_cr_arp_d_viewopt_sr?ie=UTF8&filterBy=reviewer_rated&pageNumber=3"

sessionProd1 <- bow(link,
  user_agent = "For Educational Purpose")

scrapeNodes <- function(selector){
  scrape(sessionProd1) %>%
    html_nodes(selector) %>%

```

```

    html_text(trim = TRUE)
  }

  scrapedprodName <- scrapeNodes("h1.a-size-large.a-text-ellipsis")

  scrapedReviewer <- scrapeNodes("span.a-profile-name")
  scrapedReviewer <- scrapedReviewer[1:10]

  scrapedReview <- scrapeNodes("span.a-size-base.review-text.review-text-content")
  scrapedReview <- scrapedReview[1:10]

  scrapedDate <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
  scrapedDate <- scrapedDate[1:10]

  scrapedprodRating <- scrapeNodes("span.a-icon-alt")
  scrapedprodRating <- scrapedprodRating[1:10]

  product8 <- rbind(product8, data.frame(category = scrapedCategory, name = scrapedprodName,
    reviewer = scrapedReviewer,
    review = scrapedReview,
    date = scrapedDate,
    ratings = scrapedprodRating))

  Sys.sleep(5)
#####

  link2 = "https://www.amazon.co.uk/product-reviews/B099FJCQSF/ref=cm_cr_arp_d_paging_btm_next_2?ie="

  sessionProd2 <- bow(link2,
    user_agent = "For Educational Purpose")

  scrapedReviewer2 <- scrapeNodes("span.a-profile-name")
  scrapedReviewer2 <- scrapedReviewer2[1:10]

  scrapedReview2 <- scrapeNodes("span.a-size-base.review-text.review-text-content")
  scrapedReview2 <- scrapedReview2[1:10]

  scrapedDate2 <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
  scrapedDate2 <- scrapedDate2[1:10]

  scrapedprodRating2 <- scrapeNodes("span.a-icon-alt")
  scrapedprodRating2 <- scrapedprodRating2[1:10]

  product8 <- rbind(product8, data.frame(category = scrapedCategory ,name = scrapedprodName,

```

```

        reviewer = scrapedReviewer2,
        review = scrapedReview2,
        date = scrapedDate2,
        ratings = scrapedprodRating2))

Sys.sleep(5)

#####

link3 = "https://www.amazon.co.uk/product-reviews/B099FJCQSF/ref=cm_cr_getr_d_paging_btm_next_3?i

sessionProd3 <- bow(link3,
                    user_agent = "For Educational Purpose")

scrapedReviewer3 <- scrapeNodes("span.a-profile-name")
scrapedReviewer3 <- scrapedReviewer3[1:10]

scrapedReview3 <- scrapeNodes("span.a-size-base.review-text.review-text-content")
scrapedReview3 <- scrapedReview3[1:10]

scrapedDate3 <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
scrapedDate3 <- scrapedDate3[1:10]

scrapedprodRating3 <- scrapeNodes("span.a-icon-alt")
scrapedprodRating3 <- scrapedprodRating3[1:10]

product8 <- rbind(product8, data.frame(category = scrapedCategory ,name = scrapedprodName,
                                       reviewer = scrapedReviewer3,
                                       review = scrapedReview3,
                                       date = scrapedDate3,
                                       ratings = scrapedprodRating3))

Sys.sleep(5)

#View(product8)
#write.csv(product8,"product8.csv")

product9

library(dplyr)
library(rvest)
library(polite)
library(httr)
library(selectr)

product9 = data.frame()
scrapedCategory <- "Socks"

#####

```

```

link = "https://www.amazon.co.uk/product-reviews/B07F2N8DF5/ref=cm_cr_arp_d_viewopt_sr?ie=UTF8&filter

sessionProd1 <- bow(link,
  user_agent = "For Educational Purpose")

scrapeNodes <- function(selector){
  scrape(sessionProd1) %>%
    html_nodes(selector) %>%
    html_text(trim = TRUE)
}

scrapedprodName <- scrapeNodes("h1.a-size-large.a-text-ellipsis")

scrapedReviewer <- scrapeNodes("span.a-profile-name")
scrapedReviewer <- scrapedReviewer[1:10]

scrapedReview <- scrapeNodes("span.a-size-base.review-text.review-text-content")
scrapedReview <- scrapedReview[1:10]

scrapedDate <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
scrapedDate <- scrapedDate[1:10]

scrapedprodRating <- scrapeNodes("span.a-icon-alt")
scrapedprodRating <- scrapedprodRating[1:10]

product9 <- rbind(product9, data.frame(category = scrapedCategory, name = scrapedprodName,
  reviewer = scrapedReviewer,
  review = scrapedReview,
  date = scrapedDate,
  ratings = scrapedprodRating))

Sys.sleep(5)
#####

link2 = "https://www.amazon.co.uk/product-reviews/B07F2N8DF5/ref=cm_cr_getr_d_paging_btm_next_2?i

sessionProd2 <- bow(link2,
  user_agent = "For Educational Purpose")

scrapedReviewer2 <- scrapeNodes("span.a-profile-name")
scrapedReviewer2 <- scrapedReviewer2[1:10]

scrapedReview2 <- scrapeNodes("span.a-size-base.review-text.review-text-content")
scrapedReview2 <- scrapedReview2[1:10]

scrapedDate2 <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
scrapedDate2 <- scrapedDate2[1:10]

```

```

scrapedprodRating2 <- scrapeNodes("span.a-icon-alt")
scrapedprodRating2 <- scrapedprodRating2[1:10]

product9 <- rbind(product9, data.frame(category = scrapedCategory ,name = scrapedprodName,
  reviewer = scrapedReviewer2,
  review = scrapedReview2,
  date = scrapedDate2,
  ratings = scrapedprodRating2))

Sys.sleep(5)
#####

link3 = "https://www.amazon.co.uk/product-reviews/B07F2N8DF5/ref=cm_cr_getr_d_paging_btm_next_3?i

sessionProd3 <- bow(link3,
  user_agent = "For Educational Purpose")

scrapedReviewer3 <- scrapeNodes("span.a-profile-name")
scrapedReviewer3 <- scrapedReviewer3[1:10]

scrapedReview3 <- scrapeNodes("span.a-size-base.review-text.review-text-content")
scrapedReview3 <- scrapedReview3[1:10]

scrapedDate3 <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
scrapedDate3 <- scrapedDate3[1:10]

scrapedprodRating3 <- scrapeNodes("span.a-icon-alt")
scrapedprodRating3 <- scrapedprodRating3[1:10]

product9 <- rbind(product9, data.frame(category = scrapedCategory ,name = scrapedprodName,
  reviewer = scrapedReviewer3,
  review = scrapedReview3,
  date = scrapedDate3,
  ratings = scrapedprodRating3))

Sys.sleep(5)
#View(product9)
#write.csv(product9, "product9.csv")

product10
library(dplyr)
library(rvest)
library(polite)
library(httr)
library(selectr)

```



```

product10 = data.frame()
scrapedCategory <- "Socks"

#####

link = "https://www.amazon.co.uk/product-reviews/B07R1ZCKVZ/ref=cm_cr_arp_d_viewopt_sr?ie=UTF8&filter

sessionProd1 <- bow(link,
  user_agent = "For Educational Purpose")

scrapeNodes <- function(selector){
  scrape(sessionProd1) %>%
    html_nodes(selector) %>%
    html_text(trim = TRUE)
}

scrapedprodName <- scrapeNodes("h1.a-size-large.a-text-ellipsis")

scrapedReviewer <- scrapeNodes("span.a-profile-name")
scrapedReviewer <- scrapedReviewer[1:10]

scrapedReview <- scrapeNodes("span.a-size-base.review-text.review-text-content")
scrapedReview <- scrapedReview[1:10]

scrapedDate <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
scrapedDate <- scrapedDate[1:10]

scrapedprodRating <- scrapeNodes("span.a-icon-alt")
scrapedprodRating <- scrapedprodRating[1:10]

product10 <- rbind(product10, data.frame(category = scrapedCategory, name = scrapedprodName,
  reviewer = scrapedReviewer,
  review = scrapedReview,
  date = scrapedDate,
  ratings = scrapedprodRating))

Sys.sleep(5)
#####

link2 = "https://www.amazon.co.uk/product-reviews/B07R1ZCKVZ/ref=cm_cr_getr_d_paging_btm_next_2?i

sessionProd2 <- bow(link2,
  user_agent = "For Educational Purpose")

scrapedReviewer2 <- scrapeNodes("span.a-profile-name")
scrapedReviewer2 <- scrapedReviewer2[1:10]

```

```

scrapedReview2 <- scrapeNodes("span.a-size-base.review-text.review-text-content")
scrapedReview2 <- scrapedReview2[1:10]

scrapedDate2 <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
scrapedDate2 <- scrapedDate2[1:10]

scrapedprodRating2 <- scrapeNodes("span.a-icon-alt")
scrapedprodRating2 <- scrapedprodRating2[1:10]

product10 <- rbind(product10, data.frame(category = scrapedCategory ,name = scrapedprodName,
    reviewer = scrapedReviewer2,
    review = scrapedReview2,
    date = scrapedDate2,
    ratings = scrapedprodRating2))

Sys.sleep(5)
#####

link3 = "https://www.amazon.co.uk/product-reviews/B07R1ZCKVZ/ref=cm_cr_getr_d_paging_btm_next_3?i

sessionProd3 <- bow(link3,
    user_agent = "For Educational Purpose")

scrapedReviewer3 <- scrapeNodes("span.a-profile-name")
scrapedReviewer3 <- scrapedReviewer3[1:10]

scrapedReview3 <- scrapeNodes("span.a-size-base.review-text.review-text-content")
scrapedReview3 <- scrapedReview3[1:10]

scrapedDate3 <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
scrapedDate3 <- scrapedDate3[1:10]

scrapedprodRating3 <- scrapeNodes("span.a-icon-alt")
scrapedprodRating3 <- scrapedprodRating3[1:10]

product10 <- rbind(product10, data.frame(category = scrapedCategory ,name = scrapedprodName,
    reviewer = scrapedReviewer3,
    review = scrapedReview3,
    date = scrapedDate3,
    ratings = scrapedprodRating3))

Sys.sleep(5)
#View(product10)
#write.csv(product10, "product10.csv")

```

```
All_Reviews  
  
All_Reviews <- rbind(All_Reviews)  
  
#View(All_Reviews)  
  
#write.csv(All_Reviews, "All Reviews.csv")
```