

lab_exercise#1_Olivo

Olivo, Mel Adry

2024-02-07

SCRAPING FOR NOTEBOOKS

```
#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)

notebook_products <- data.frame()

notebooklink1 = "https://www.amazon.co.uk/s?k=notebook&crid=3PE2BSTUHI0D6&qid=1707353799&srefix=notebo

session1 <- bow(notebooklink1,
               user_agent = "Educational Purpose")

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

scrapedCategory <- "Notebook"

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]

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

#####

notebooklink2 = "https://www.amazon.co.uk/s?k=notebook&page=2&crd=3PE2BSTUHI0D6&qid=1707353804&sprefix:

session2 <- bow(notebooklink2,
  user_agent = "For Educational Purpose")

scrapedCategory2 <- "Notebook"

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]

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

#####

notebooklink3 = "https://www.amazon.co.uk/s?k=notebook&page=3&crd=3PE2BSTUHI0D6&qid=1707354306&sprefix:

session3 <- bow(notebooklink3,
  user_agent = "For Educational Purpose")

scrapedCategory3 <- "Notebook"

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]

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

notebook_products <- notebook_products[1:100,]

#View(notebook_products)

#####

```

SCRAPING FOR PENCILS

```

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

pencil_products <- data.frame()

pencillink1_1_1 = "https://www.amazon.co.uk/s?k=pencil&crd=ROFYFPYBSIDI&qid=1707354893&srefix=penicl%"

session4_4_4 <- bow(pencillink1_1_1,
  user_agent = "Education")

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

scrapedCategory4 <- "Pencil"

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

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

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

```

```

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

pencil_products <- rbind(pencil_products, data.frame(category = scrapedCategory4, name = scrapedpencilName,
  price = scrapedpencilPrice,
  ratings = scrapedpencilRatings,
  no_of_reviews = scrapedpencilReviews))

#####

pencillink2_2_2 = "https://www.amazon.co.uk/s?k=pencil&crd=R0FYFPYBSIDI&qid=1707354893&srefix=penicl%"

session5_5_5 <- bow(pencillink2_2_2,
  user_agent = "Educational Purpose")

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

scrapedCategory5 <- "Pencil"

scrapedpencilName2 <- scrapeNodes("h2.a-size-mini.a-spacing-none.a-color-base.s-line-clamp-4")
scrapedpencilName2 <- scrapedpencilName2[1:45]

scrapedpencilPrice2 <- scrapeNodes("span.a-offscreen")
scrapedpencilPrice2 <- scrapedpencilPrice2[1:45]

scrapedpencilRatings2 <- scrapeNodes("span.a-icon-alt")
scrapedpencilRatings2 <- scrapedpencilRatings2[1:45]

scrapedpencilReviews2 <- scrapeNodes("span.a-size-base.s-underline-text")
scrapedpencilReviews2 <- scrapedpencilReviews2[1:45]

pencil_products <- rbind(pencil_products, data.frame(category = scrapedCategory5, name = scrapedpencilName,
  price = scrapedpencilPrice2,
  ratings = scrapedpencilRatings2,
  no_of_reviews = scrapedpencilReviews2))

#####

pencillink3_3_3 = "https://www.amazon.co.uk/s?k=pencil&crd=R0FYFPYBSIDI&qid=1707354893&srefix=penicl%"

session6_6_6 <- bow(pencillink3_3_3,
  user_agent = "Educational Purpose")

scrapeNodes <- function(selector){

```

```

scrape(session6_6_6) %>%
  html_nodes(selector) %>%
  html_text(trim = TRUE)
}

scrapedCategory6 <- "Pencil"

scrapedpencilName3 <- scrapeNodes("h2.a-size-mini.a-spacing-none.a-color-base.s-line-clamp-4")
scrapedpencilName3 <- scrapedpencilName3[1:45]

scrapedpencilPrice3 <- scrapeNodes("span.a-offscreen")
scrapedpencilPrice3 <- scrapedpencilPrice3[1:45]

scrapedpencilRatings3 <- scrapeNodes("span.a-icon-alt")
scrapedpencilRatings3 <- scrapedpencilRatings3[1:45]

scrapedpencilReviews3 <- scrapeNodes("span.a-size-base.s-underline-text")
scrapedpencilReviews3 <- scrapedpencilReviews3[1:45]

pencil_products <- rbind(pencil_products, data.frame(category = scrapedCategory5, name = scrapedpencilName3,
  price = scrapedpencilPrice2,
  ratings = scrapedpencilRatings2,
  no_of_reviews = scrapedpencilReviews2))

pencil_products <- pencil_products[1:100,]

#View(pencil_products)

```

MERGING TWO PRODUCTS

```

scraped_products <- rbind(notebook_products, pencil_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 <- "Notebook"

###

link = "https://www.amazon.co.uk/Blue-Acorn-Spiral-Bound-Sketchbook/product-reviews/BOCH3KXXSN/ref=cm

```

```

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/Blue-Acorn-Spiral-Bound-Sketchbook/product-reviews/BOCH3KXXSN/r"

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/Blue-Acorn-Spiral-Bound-Sketchbook/product-reviews/BOCH3KXXSN/r"

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)

product2

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

product2 = data.frame()
scrapedCategory <- "Notebook"

```

```
###
```

```
link = "https://www.amazon.co.uk/Oxford-Notebook-Subject-College-63756/product-reviews/B08VF9B772/ref="
```

```
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/Oxford-Notebook-Subject-College-63756/product-reviews/B08VF9B772/ref="
```

```
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/Oxford-Notebook-Subject-College-63756/product-reviews/B08VF9B777"

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)

product3
library(dplyr)
library(rvest)
library(polite)
library(httr)
library(selectr)

```

```

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

###

link = "https://www.amazon.co.uk/Leitz-Covered-Notebook-Squared-Sheets/product-reviews/B073QYZD5Q/ref=

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/Leitz-Covered-Notebook-Squared-Sheets/product-reviews/B073QYZD5Q/ref=

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

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

```

```

scrappedReview2 <- scrapeNodes("span.a-size-base.review-text.review-text-content")
scrappedReview2 <- scrappedReview2[1:10]

scrappedDate2 <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
scrappedDate2 <- scrappedDate2[1:10]

scrappedprodRating2 <- scrapeNodes("span.a-icon-alt")
scrappedprodRating2 <- scrappedprodRating2[1:10]

product3 <- rbind(product3, data.frame(category = scrappedCategory ,name = scrappedprodName,
  reviewer = scrappedReviewer2,
  review = scrappedReview2,
  date = scrappedDate2,
  ratings = scrappedprodRating2))

Sys.sleep(5)
####

link3 = "https://www.amazon.co.uk/Leitz-Covered-Notebook-Squared-Sheets/product-reviews/B073QYZD5"

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

scrappedReviewer3 <- scrapeNodes("span.a-profile-name")
scrappedReviewer3 <- scrappedReviewer3[1:10]

scrappedReview3 <- scrapeNodes("span.a-size-base.review-text.review-text-content")
scrappedReview3 <- scrappedReview3[1:10]

scrappedDate3 <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")
scrappedDate3 <- scrappedDate3[1:10]

scrappedprodRating3 <- scrapeNodes("span.a-icon-alt")
scrappedprodRating3 <- scrappedprodRating3[1:10]

product3 <- rbind(product3, data.frame(category = scrappedCategory ,name = scrappedprodName,
  reviewer = scrappedReviewer3,
  review = scrappedReview3,
  date = scrappedDate3,
  ratings = scrappedprodRating3))

Sys.sleep(5)
#View(product3)

```

product4

```

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

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

###

link = "https://www.amazon.co.uk/Oxford-My-Notes-Wirebound-Perforated/product-reviews/B00BI5N0GW/ref=

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/Oxford-My-Notes-Wirebound-Perforated/product-reviews/B00BI5N0GW"

```

```

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/Oxford-My-Notes-Wirebound-Perforated/product-reviews/B00BI5NOGW"

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)

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

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

###

link = "https://www.amazon.co.uk/Amazon-Basics-Casebound-Notebook-21x13-3x1-6/product-reviews/B01DN8T

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/Amazon-Basics-Casebound-Notebook-21x13-3x1-6/product-reviews/B0
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/Amazon-Basics-Casebound-Notebook-21x13-3x1-6/product-reviews/B0
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)

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

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

###

link = "https://www.amazon.co.uk/STAEDTLER-121-2-BK5D-Noris-pencils/product-reviews/B093L3F7GN/ref=cm

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/STAEDTLER-121-2-BK5D-Noris-pencils/product-reviews/B093L3F7GN/r"

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/STAEDTLER-121-2-BK5D-Noris-pencils/product-reviews/B093L3F7GN/r"

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)

product7
library(dplyr)
library(rvest)
library(polite)
library(httr)
library(selectr)

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

###

link = "https://www.amazon.co.uk/STAEDTLER-121-2-BK5D-Noris-pencils/product-reviews/B093L3F7GN/ref=cm

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/STAEDTLER-121-2-BK5D-Noris-pencils/product-reviews/B093L3F7GN/r"

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/STAEDTLER-121-2-BK5D-Noris-pencils/product-reviews/B093L3F7GN/r"

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)

product8
library(dplyr)
library(rvest)
library(polite)
library(httr)
library(selectr)

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

###

link = "https://www.amazon.co.uk/AmazonBasics-Presharpened-Wood-Cased-Pencils/product-reviews/B071JM6"

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/AmazonBasics-Presharpened-Wood-Cased-Pencils/product-reviews/B0"

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/AmazonBasics-Presharpened-Wood-Cased-Pencils/product-reviews/B00
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)

product9
library(dplyr)
library(rvest)
library(polite)
library(httr)
library(selectr)

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

###

link = "https://www.amazon.co.uk/Ecolutions-Evolution-655%C3%82-HB-Pencil-Eraser/product-reviews/B0000

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/Ecolutions-Evolution-655%C3%82-HB-Pencil-Eraser/product-reviews"

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/Ecolutions-Evolution-655%C3%82-HB-Pencil-Eraser/product-reviews,

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)

product10

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

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

###

link = "https://www.amazon.co.uk/BIC-Evolution-Original-Graphite-Pencils/product-reviews/B003V8BIFE/r

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/BIC-Evolution-Original-Graphite-Pencils/product-reviews/B003V8B"

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/BIC-Evolution-Original-Graphite-Pencils/product-reviews/B003V8B"

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)

all_reviews <- rbind(product1, product2, product3, product4, product5, product6, product7, product8, pr
#all_reviews

```