LabExer2

Rey Angelo Calopez

2024-02-08

Scraping 50 Amazon Products

```r
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
```

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

```
##
## Attaching package: 'readr'

## The following object is masked from 'package:rvest':
##
##     guess_encoding
```

```r
scrapeAmazonReviews <- function(product_name, url_stub, pages) {
  product_df <- data.frame()

  for (page in pages) {
    link <- paste0("https://www.amazon.com.au/",url_stub,"/ref=cm_cr_arp_d_paging_btm_next_", page, "?id
    session <- bow(link, user_agent = "Educational Purpose")

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

    scrapedTitle <- scrapeNodes("span.a-size-base.review-title.a-color-base.review-title-content.a-text-
    scrapedReviewer <- scrapeNodes("span.a-profile-name")[1:10]
    scrapedReview <- scrapeNodes("span.a-size-base.review-text.review-text-content")[1:10]
    scrapedDate <- scrapeNodes("span.a-size-base.a-color-secondary.review-date")[1:10]
    scrapedRating <- scrapeNodes("span.a-icon-alt")[1:10]
```

```r
    scrapedType <- scrapeNodes("span.a-size-mini.a-color-state.a-text-bold")[1:10]

    product_df <- rbind(product_df, data.frame(
      prod_name = product_name,
      title = scrapedTitle,
      reviewer = scrapedReviewer,
      review = scrapedReview,
      date = scrapedDate,
      ratings = scrapedRating,
      type_of_purchase = scrapedType
    ))

    #Sys.sleep(1)
  }

  return(product_df)
}
```

1-10 Slides: Crocs

```r
product1 <- scrapeAmazonReviews(
  "Crocs Unisex-Adult Classic Cozzzy Fuzzy Sandals Slide",
  "Crocs-Unisex-Adult-Classic-Platform-Slippers/product-reviews/B0BFNN9TPR",
  3:7
)


product2 <- scrapeAmazonReviews(
  "Crocs Unisex Adult Mellow Recovery Slide",
  "Crocs-Unisex-Adult-Mellow-Recovery/product-reviews/B0BFNN6P1N",
  6:10
)


product3 <- scrapeAmazonReviews(
  "Crocs Unisex Adults Bayaband Slide",
  "Crocs-Womens-Unisex-Adult-Bayaband-Sandal/product-reviews/B07B8G7KRP",
  3:7
)


product4 <- scrapeAmazonReviews(
  "Crocs Unisex Classic Sandal",
  "Crocs-Classic-Sandal-Unisex-Slide/product-reviews/B088YT7JS4",
  3:7
)


product5 <- scrapeAmazonReviews(
  "Skechers Women's On-The-go 600-16259 Slide Sandal",
  "Skechers-Womens-Open-Sandals-Textile/product-reviews/B07TYVQRM1",
```

```
    3:7
)

product6 <- scrapeAmazonReviews(
  "Crocs womens Classic Platform Slide",
  "Crocs-Womens-Classic-Platform-Sandals/product-reviews/B0B4SGW2FB",
  3:7
)


product7 <- scrapeAmazonReviews(
  "Crocs Unisex Kids Classic Pokemon Clog T",
  "Crocs-Unisex-Classic-Pokemon-White/product-reviews/B09P1J5KTN",
  3:7
)


product8 <- scrapeAmazonReviews(
  "Crocs Women's Splash Slides Sandal",
  "Crocs-Womens-Splash-Slides-Sandal/product-reviews/B0B4SGCP2K",
  3:7
)


product9 <- scrapeAmazonReviews(
  "Crocs Unisex-Adult Baya Ii Slides Sandal",
  "Crocs-Unisex-Adult-Slides-Sandal-Black/product-reviews/B0B4SDYCF8",
  3:7
)


product10 <- scrapeAmazonReviews(
  "Crocs Men's Clogs",
  "Crocs-Unisex-Literide-Sandal-Stucco/product-reviews/B0C5K141PK",
  6:10
)
```

11-20 Slides: Adidas

```
product11 <- scrapeAmazonReviews(
  "adidas Men's Adilette Aqua Slipper",
  "adidas-Adilette-Aqua-AD-Black-Sandals/product-reviews/B07F73JQ96",
  3:7
)


product12 <- scrapeAmazonReviews(
  "adidas Unisex-Adult Adissage Slides Sandal",
  "adidas-Unisex-Adilette-3-Stripes-Slippers/product-reviews/B09HZDJN4X",
  3:7
)


product13 <- scrapeAmazonReviews(
  "adidas Men's Adilette Aqua 3-Stripes Slides Sandal",
```

```r
  "adidas-Adilette-Sandal-Footwear-Metallic/product-reviews/B07ZQCNRPS",
  3:7
)


product14 <- scrapeAmazonReviews(
  "adidas",
  "adidas-Performance-Adissage-Black-Sandals/product-reviews/B07J4P9XSN",
  3:7
)


product15 <- scrapeAmazonReviews(
  "adidas Women's Adilette",
  "adidas-G28695-Womens-Adilette-Sandal/product-reviews/B07D9FVM29",
  3:7
)


product16 <- scrapeAmazonReviews(
  "adidas Men's Adilette Comfort Slide Sandals, Core Black/Cloud White/Cloud White, 4 US",
  "adidas-Unisex-Child-Boys-Girls-Adilette/product-reviews/B07SDTSNLK",
  3:7
)


product17 <- scrapeAmazonReviews(
  "adidas Women's Adilette Clogs Slide Sandal",
  "Adidas-Adilette-Comfort-Black-Sandals/product-reviews/B073RK4BPP",
  3:7
)


product18 <- scrapeAmazonReviews(
  "adidas Originals Adilette Slip-On Slides",
  "adidas-Unisexs-Adilette-Slide-Sandal/product-reviews/B0BZ7G4RZD",
  3:7
)


product19 <- scrapeAmazonReviews(
  "adidas",
  "Adidas-Originals-Adilette-Sport-Slide/product-reviews/B002W6ZAY6",
  4:8
)


product20 <- scrapeAmazonReviews(
  "adidas Unisex-Adult Adissage Sandal",
  "adidas-Performance-Adissage-Black-Sandals/product-reviews/B07K2KHYZS",
  3:7
)
```

Scraping 21-30 Nike Products

```r
product21 <- scrapeAmazonReviews(
  "Nike Men's Victori One Slide",
  "Nike-Victori-Slide-Running-Black-White/product-reviews/B08Q9MLZZM",
  3:7
)


product22 <- scrapeAmazonReviews(
  "Nike Men's Revolution 6 Nn Sneaker, Black White Iron Grey, 11.5 UK, Black White Iron Grey, 12.5 US",
  "Nike-Revolution-Flyease-Running-Black/product-reviews/B098PCGH4F",
  3:7
)


product23 <- scrapeAmazonReviews(
  "Nike Men's Victori One Shower Slide Trail Running Shoe, Black/White-Black",
  "Shower-Slide-Beach-Shoes-White/product-reviews/B08QBM6GKF",
  3:7
)


product24 <- scrapeAmazonReviews(
  "Nike Damen Victori One Slide Sandal",
  "Nike-Womens-Victori-Running-White-Black/product-reviews/B08P5LMTG8",
  3:7
)


product25 <- scrapeAmazonReviews(
  "Nike Men's Victori One Slide Print Slippers",
  "Nike-Shower-Slide-Beach-Shoes/product-reviews/B08QBJB2Q5",
  3:7
)

product26 <- scrapeAmazonReviews(
  "Nike Unisex Everyday Cushion Ankle Training Socks (3 Pair)",
  "Nike-Unisex-Everyday-Cushion-Training/product-reviews/B07GBH42SG",
  3:7
)


product27 <- scrapeAmazonReviews(
  "Nike Women's W VICTORI ONE Slide Gymnastics Shoe",
  "Nike-womens-Slider-Black-9/product-reviews/B08NFC2MPJ",
  3:7
)


product28 <- scrapeAmazonReviews(
  "Nike Kawa Slide Se (TD), Unisex-Children and Youth Sneakers",
  "Nike-Slide-Infant-Toddler-White/product-reviews/B086R7TBSK",
  3:7
)
```

```r
product29 <- scrapeAmazonReviews(
  "Nike Everyday Cushion Crew Training Socks",
  "Unisex-Everyday-Cushion-Black-X-Large/product-reviews/B07FKF3R3F",
  3:7
)


product30 <- scrapeAmazonReviews(
  "Nike Men's M NK DRY PARK VII JSY SS T-shirt",
  "Nike-Mens-T-Shirt-University-White/product-reviews/B07W4SY5QH",
  3:7
)
```

31-40 Under Armour Products

```r
product31 <- scrapeAmazonReviews(
  "Under Armour Men's Locker Camo Slide Sandal",
  "Under-Armour-Walking-Black-White/product-reviews/B08FTKT8S6",
  3:7
)


product32 <- scrapeAmazonReviews(
  "Under Armour Men's Tech Short Sleeve T-Shirt",
  "Under-Armour-Sleeve-Graphite-Medium/product-reviews/B0785VMVF4",
  3:7
)


product33 <- scrapeAmazonReviews(
  "Under Armour Men's UA Locker IV Slides Sandal",
  "Under-Armour-Mens-Locker-Gray/product-reviews/B087T3J1DR",
  3:7
)


product34 <- scrapeAmazonReviews(
  "Under Armour Men's UA Locker IV Slides Sandal",
  "Under-Armour-Locker-Slides-Academy/product-reviews/B093Y7SNVG",
  3:7
)


product35 <- scrapeAmazonReviews(
  "Under Armour Men's Ansa Fix Slide Sandal",
  "Under-Armour-Slide-Sandal-Royal/product-reviews/B082JH3P5W",
  3:7
)


product36 <- scrapeAmazonReviews(
  "Under Armour",
  "Under-Armour-HeatGear-Shorts-Black/product-reviews/B089T92G55",
```

```
   3:7
)


product37 <- scrapeAmazonReviews(
  "Under Armour",
  "Under-Armour-Unisex-3021286-400-8-Academia/product-reviews/B07KTMZTYB",
  3:7
)


product38 <- scrapeAmazonReviews(
  "Under Armour Unisex Adults Ua Core Pth Sl Beach & Pool Shoes",
  "Under-Armour-Unisex-Adults-Beach/product-reviews/B07PWNP4C2",
  3:7
)


product39 <- scrapeAmazonReviews(
  "Under Armour Men's UA Surge 3",
  "Under-Armour-Surge-Running-Shoes/product-reviews/B09691RG2X",
  3:7
)


product40 <- scrapeAmazonReviews(
  "Under Armour",
  "Under-Armour-Charged-Assert-Sneaker/product-reviews/B09XBVQNWQ",
  3:7
)
```

41-50 THE NORTH FACE Products

```
product41 <- scrapeAmazonReviews(
  "THE NORTH FACE Men's Base Camp Iii Flip-Flop",
  "North-Face-Slide-White-Black/product-reviews/B08NWFCD7R",
  3:7
)


product42 <- scrapeAmazonReviews(
  "THE NORTH FACE Unisex Vault Rucksack",
  "North-Face-Vault-TNF-Black/product-reviews/B0823VZ3PH",
  3:7
)


product43 <- scrapeAmazonReviews(
  "THE NORTH FACE Unisex's Jester Shoulder Bag, One Size, Tnf Black",
  "NORTH-FACE-Unisexs-Jester-Shoulder/product-reviews/B09JQKX6MG",
  3:7
)
```

```r
product44 <- scrapeAmazonReviews(
  "THE NORTH FACE Women's Base Camp Iii Sandal, M",
  "North-Womens-Slide-White-Black/product-reviews/B08NWDTTFT",
  3:7
)


product45 <- scrapeAmazonReviews(
  "The North Face Women's Thermoball Traction Mule",
  "North-Face-Thermoball-Traction-Slippers/product-reviews/B07N98SS39",
  3:7
)


product46 <- scrapeAmazonReviews(
  "THE NORTH FACE Women's Etip Gloves",
  "North-Womens-Recycled-Glove-Black/product-reviews/B083LWD6LT",
  3:7
)


product47 <- scrapeAmazonReviews(
  "THE NORTH FACE Jester Backpack",
  "North-Face-Womens-Jester-Black/product-reviews/B0823VNN1P",
  3:7
)


product48 <- scrapeAmazonReviews(
  "Tzowla Travel Laptop Backpack,Durable Water Resistant Anti-Theft 50L Extra Large 17.3 Inch with USB (
  "Business-Backpack-Resistant-Anti-theft-Backpacks/product-reviews/B075SZQN9Q",
  3:7
)


product49 <- scrapeAmazonReviews(
  "THE NORTH FACE Men's Jim Beanie",
  "North-Face-Beanie-Shady-Blue/product-reviews/B09TPVKTMX",
  3:7
)


product50 <- scrapeAmazonReviews(
  "THE NORTH FACE Men's Bones Recycled Beanies",
  "North-Bones-Recycled-Beanie-Cordovan/product-reviews/B09TPVQR4G",
  3:7
)

allProds = data.frame();

allProds <- rbind(allProds, product1, product2, product3, product4, product5, product6, product7, produ
                  product11, product12, product13, product14, product15, product16, product17, product18,
                  product21, product22, product23, product24, product25, product26, product27, product28,
                  product31, product32, product33, product34, product35, product36, product37, product38,
```

```
                product41, product42, product43, product44, product45, product46, product47, product48,

#View(allProds)

write.csv(allProds, "/cloud/project/LabExer#2/ScrapedAmazonProducts.csv")
```