RWorksheet#5_group.Rmd

Baylon_Calvario_Calzado

2024-11-01

```
library(polite)
library(kableExtra)
library(rmarkdown)
library(httr)
library(rvest)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following object is masked from 'package:kableExtra':
##
##     group_rows
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(stringr)
```

```
urls <- c('https://www.amazon.com/s?k=skin+care&crid=3FODLJK4URUE9&sprefix=skin%2Caps%2C510&ref=nb_sb_s

nameofproduct <- vector("list", length(urls))
price <- vector("list", length(urls))
description <- vector("list", length(urls))
Ratings <- vector("list", length(urls))
Reviews <- vector("list", length(urls))
```

```
df <- list()

for (i in seq_along(urls)) {
    session <- bow(urls[i], user_agent = "Educational")
    webpage <- scrape(session)

nameofproduct[[i]] <-  webpage %>%
  html_nodes('.a-size-base-plus.a-color-base') %>%
  html_text() %>%
  head(30)

price[[i]] <-  webpage %>%
  html_nodes('.a-price-whole') %>%
```

```
  html_text() %>%
  head(30)

description[[i]] <-  webpage %>%
  html_nodes('.a-size-mini.a-spacing-none.a-color-base s-line-clamp-3') %>%
  html_text() %>%
  head(30)

Ratings[[i]] <-  webpage %>%
  html_nodes('.a-icon.a-icon-star-small.a-star-small-4-5') %>%
  html_text() %>%
  head(30)

Reviews[[i]] <-  webpage %>%
  html_nodes('.a-size-base-plus.a-color-base') %>%
  html_text() %>%
  head(30)
}
```

```
skincare <- data.frame(nameofproduct[[1]],
                       Price = price[[1]],
                       Ratings = Ratings[[1]],
                       stringsAsFactors = FALSE
                       )
head(skincare)
```

```
##
## 1
## 2 Moisturizing Cream | Body and Face Moisturizer for Dry Skin | Body Cream with Hyaluronic Acid and
## 3
## 4                                    Toleriane Double Repair Face Moisturizer | Daily Moisturizer Fa
## 5
## 6
##   Price            Ratings
## 1   17. 4.7 out of 5 stars
## 2   23. 4.6 out of 5 stars
## 3   19. 4.5 out of 5 stars
## 4    9. 4.4 out of 5 stars
## 5   15. 4.7 out of 5 stars
## 6   35. 4.5 out of 5 stars
```

```
perfume <- data.frame(nameofproduct[[2]],
                       Price = price[[2]],
                       Ratings = Ratings[[2]],
                       stringsAsFactors = FALSE
                       )
head(perfume)
```

```
##
## 1
## 2                                                                                          Lattafa
## 3 Billie Eilish Eau de Parfum Spray Perfume for Women, Notes of Sugared Petals, Vanilla & Warm Musk,
## 4                                                                                        Lattafa Per
## 5                                                                                  Victoria's Secre
## 6                                                                                  Versace Bright Cr
```

```
##   Price            Ratings
## 1   21. 4.4 out of 5 stars
## 2   52. 4.6 out of 5 stars
## 3   27. 4.5 out of 5 stars
## 4   16. 4.6 out of 5 stars
## 5   31. 4.6 out of 5 stars
## 6   23. 4.4 out of 5 stars
```

```
bags <- data.frame(nameofproduct[[3]],
                       Price = price[[3]],
                       Ratings = Ratings[[3]],
                       stringsAsFactors = FALSE
                       )
head(bags)
```

```
##
## 1
## 2                                                                    Large Canvas Tote Bag for Women
## 3
## 4
## 5 Travel Backpack for Women, Carry On Backpack with USB Charging Port & Shoe Pouch, TSA 15.6inch Lap
## 6
##   Price            Ratings
## 1   21. 4.4 out of 5 stars
## 2   23. 4.7 out of 5 stars
## 3    8. 4.5 out of 5 stars
## 4   34. 4.7 out of 5 stars
## 5   61. 4.7 out of 5 stars
## 6   59. 4.3 out of 5 stars
```

```
shoes <- data.frame(nameofproduct[[4]],
                       Price = price[[4]],
                       Ratings = Ratings[[4]],
                       stringsAsFactors = FALSE
                       )
head(shoes)
```

```
##              nameofproduct..4.. Price            Ratings
## 1              Amazon Essentials   42. 4.4 out of 5 stars
## 2 Women's Tall Block Heel Boots   64. 4.5 out of 5 stars
## 3 Amazon's Choice: Overall Pick   49. 4.7 out of 5 stars
## 4                  Under Armour   89. 4.6 out of 5 stars
## 5       Men's Charged Assert 10   64. 4.6 out of 5 stars
## 6                         adidas   46. 4.3 out of 5 stars
```

```
pants <- data.frame(nameofproduct[[5]],
                       Price = price[[5]],
                       Ratings = Ratings[[5]],
                       stringsAsFactors = FALSE
                       )
head(pants)
```

```
##                                                  nameofproduct..5..
## 1                                 Amazon's Choice: Overall Pick
## 2                                           Wrangler Authentics
## 3                              Men's Relaxed Fit Stretch Cargo Pant
```

```
## 4                                                                  Ultra Performance
## 5 3 Pack Fleece Active Tech Joggers for Men, Mens Sweatpants with Zipper Pockets
## 6                                                                  Fruit of the Loom
##   Price            Ratings
## 1   25. 4.5 out of 5 stars
## 2   35. 4.4 out of 5 stars
## 3   12. 4.4 out of 5 stars
## 4   21. 4.3 out of 5 stars
## 5   19. 4.5 out of 5 stars
## 6    3. 4.3 out of 5 stars
```

#ffdgbfh