Group 4

Miko Pama, Rey-Ann Laguda, Neo Michael Kenneth Espia

2023-12-22

```r
#install.packages("rvest")

library(rvest)
url <-"https://www.amazon.com/WHITIN-Insulated-Waterproof-Anti-Slip-invierno/dp/B08CVSLG82/ref=sr_1_2_ss
url
```

```
## [1] "https://www.amazon.com/WHITIN-Insulated-Waterproof-Anti-Slip-invierno/dp/B08CVSLG82/ref=sr_1_2_s
```

```r
scrape <- read_html (url)
```

```r
scrape_amazon_reviews <- function(url, max_reviews= 300) {

scrape <- read_html(url)

username <- scrape %>%

html_nodes(".a-profile-name") %>%

html_text()

ratings <- scrape %>%

html_nodes (".review-rating") %>%

html_text()

reviews <- scrape %>%

html_nodes(".review-text-content span") %>%

html_text()

# Find the minimum length among the vectors

min_length <- min(length (username), length(ratings), length(reviews))

# Create a data frame only with complete rows

data <- data.frame(
Username = username[1:min_length],
Ratings = ratings[1:min_length],
Reviews = reviews[1:min_length]
)
```

```
return(data)
}

reviews_data <- scrape_amazon_reviews(url, max_reviews = 300)
reviews_data
```

```
##   Username Ratings Reviews
## 1     <NA>    <NA>    <NA>
```