

# RWorksheet#5a\_group6(Vicinte, Eusuya, Montevalto)

Group6

2024-11-18

```
install.packages("polite")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.4'  
## (as 'lib' is unspecified)
```

```
install.packages("dplyr")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.4'  
## (as 'lib' is unspecified)
```

```
install.packages("httr")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.4'  
## (as 'lib' is unspecified)
```

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

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.4'  
## (as 'lib' is unspecified)
```

```
library(polite)
```

```
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(httr)
```

```
library(rvest)
```

```
polite::use_manners(save_as = 'polite_scrape.R')
```

```
## v Setting active project to "/cloud/project".
```

```
## v Writing 'polite_scrape.R'.
```

```
## [ ] Edit 'polite_scrape.R'.
```

```
url <- 'https://www.amazon.com/s?k=smartphone&crd=39IXJIHXRLE3&qid=1731928902&srefix=smartphone%2Cap'
```

```
session <- bow(url,
```

```

        user_agent = "Educational")
session

## <polite session> https://www.amazon.com/s?k=smartphone&crid=39IXJIHHRLE3&qid=1731928902&sprefix=sma
##   User-agent: Educational
##   robots.txt: 137 rules are defined for 4 bots
##   Crawl delay: 5 sec
##   The path is scrapable for this user-agent

title_list <- scrape(session) %>%
  html_nodes('h2.a-size-mini.a-spacing-none.a-color-base.s-line-clamp-2') %>%
  html_text
title_list

## [1] "Moto G Stylus 5G | 2024 | Unlocked | Made for US 8/256GB | 50MP Camera | Caramel Latte  "
## [2] "SAMSUNG Galaxy S24 Ultra Cell Phone, 256GB AI Smartphone, Unlocked Android, 200MP, 100x Zoom C
## [3] "SAMSUNG Galaxy A35 5G A Series Cell Phone, 128GB Unlocked Android Smartphone, AMOLED Display,
## [4] "Tracfone | Motorola Moto g Play 2024 | Locked | 64GB | 5000mAh Battery | 50MP Quad Pixel Camera
## [5] "Moto G Play | 2024 | Unlocked | Made for US 4/64GB | 50MP Camera | Sapphire Blue  "
## [6] "Samsung Galaxy A03s Cell Phone, AT&T GSM Unlocked Android Smartphone, 32GB, Long Lasting Batter
## [7] "Samsung Galaxy A05 A055M 64GB Dual-SIM GSM Unlocked Android Smartphone (Latin America Version)
## [8] "Samsung Galaxy A15 (SM-155M/DSN), 128GB 6GB RAM, Dual SIM, Factory Unlocked GSM, International
## [9] "Samsung Galaxy A15 A155M 128GB Dual-SIM GSM Unlocked Android Smartphone (Latin America Version)
## [10] "Samsung Galaxy S21 5G, US Version, 128GB, Phantom Gray - Unlocked (Renewed)  "
## [11] "OnePlus Nord N30 5G | Unlocked Dual-SIM Android Smart Phone | 6.7\" LCD Display | 8 +128GB | 5
## [12] "SAMSUNG, Galaxy A04e (SM-A042M/DS) Dual SIM 32GB, 6.5\" GSM Unlocked, International Version-Black
## [13] "SAMSUNG Galaxy S21 FE 5G Cell Phone, Factory Unlocked Android Smartphone, 128GB, 120Hz Display
## [14] "Samsung Galaxy S21 Ultra 5G, US Version, 128GB, Phantom Black - Unlocked (Renewed)  "
## [15] "TracFone | Motorola Moto g Pure | Locked | 32GB | 4000 mAh Battery | 13 MP Dual Camera System
## [16] "Samsung Galaxy S22 Smartphone, Factory Unlocked Android Cell Phone, 128GB, 8K Camera & Video,

price <- scrape(session) %>%
  html_nodes('span.a-price-whole') %>%
  html_text
price

## [1] "249." "949." "299." "49." "109." "65." "93." "174." "159." "206."
## [11] "299." "86." "294." "279." "49." "237."

ratings <- scrape(session) %>%
  html_nodes('span.a-size-base.s-underline-text') %>%
  html_text
ratings

## [1] "106"      "3,277"    "415"      "100"      "1,498"    "753"      "200"      "1,112"
## [9] "682"      "10,704"   "1,046"    "774"      "14,934"   "2,594"    "2,857"    "2,591"

url2 <- 'https://www.amazon.com/s?k=smartphone&page=2&crid=39IXJIHHRLE3&qid=1731928557&sprefix=smartph

session2 <- bow(url2,
  user_agent = "Educational")
session2

## <polite session> https://www.amazon.com/s?k=smartphone&page=2&crid=39IXJIHHRLE3&qid=1731928557&spre
##   User-agent: Educational
##   robots.txt: 137 rules are defined for 4 bots
##   Crawl delay: 5 sec

```

```
## The path is scrapable for this user-agent
```

```
title_list2 <- scrape(session2) %>%
  html_nodes('h2.a-size-mini.a-spacing-none.a-color-base.s-line-clamp-2') %>%
  html_text
title_list2
```

```
## [1] "OnePlus Nord N200 | 5G Unlocked U.S Version | 6.49\" Full HD+LCD Screen | 90Hz Smooth Display
## [2] "OnePlus Nord N30 5G | Unlocked Dual-SIM Android Smart Phone | 6.7\" LCD Display | 8 +128GB | 5
## [3] "Google Pixel 4a with 5G - Android Phone - New Unlocked Smartphone with Night Sight and Ultrawid
## [4] "Samsung Galaxy A15 (SM-155M/DSN), 128GB 6GB RAM, Dual SIM, Factory Unlocked GSM, International
## [5] "Samsung Galaxy A14 5G A Series, Factory Unlocked, 64GB, US Version, Black (Renewed) "
## [6] "TracFone | Motorola Moto g Pure | Locked | 32GB | 4000 mAh Battery | 13 MP Dual Camera System
## [7] "Alcatel 1 (16GB) 5.0\" Full View Display, Removable Battery, FM Radio, Dual SIM GSM Unlocked U
## [8] "Samsung Galaxy A25 5G A256E 128GB Dual-SIM GSM Unlocked Android Smartphone (Latin Variant/US C
## [9] "Motorola Moto G54 Dual-SIM 256GB ROM + 8GB RAM (Only GSM | No CDMA) Factory Unlocked 5G Smartph
## [10] "Motorola Moto G Pure | 2021 | 2-Day battery | Unlocked | Made for US by Motorola | 3/32GB | 13
## [11] "Google Pixel 8 - Unlocked Android Smartphone with Advanced Pixel Camera, 24-Hour Battery, and
## [12] "Samsung Galaxy S20 FE 5G, 128GB, Cloud Navy - Unlocked (Renewed) "
## [13] "OnePlus 9 Winter Mist, 5G Unlocked Android Smartphone U.S Version, 8GB RAM+128GB Storage, 120H
## [14] "Google Pixel 9 - Unlocked Android Smartphone with Gemini, 24-Hour Battery, Advanced Camera, and
## [15] "Xiaomi Redmi Note 13 PRO 5G + 4G LTE (256GB + 8GB) 6.67\" 200MP Triple (Tmobile Mint Tello & G
## [16] "SAMSUNG Galaxy A25 5G A Series Cell Phone, 128GB Unlocked Android Smartphone, AMOLED Display, A
```

```
combined_titles <- c(title_list, title_list2)
```

```
title_df1 <- data.frame(Title = title_list, stringsAsFactors = FALSE)
title_df2 <- data.frame(Title = title_list2, stringsAsFactors = FALSE)
```

```
combined_df <- rbind(title_df1, title_df2)
combined_df
```

```
##
## 1
## 2 SAMSUNG Galaxy S24 Ultra Cell Phone, 256GB AI Smartphone,
## 3 SAMSUNG Galaxy A35 5G A Series Cell Phone, 128GB Unlocked Android Smartphone, AMOLED Dis
## 4 Tracfone | Motorola Moto g Play 202
## 5
## 6 Samsung Galaxy A03s Cell Phone, AT&T GSM Unlocked Android Smartphone
## 7 Samsung
## 8 Samsung Galaxy A15 (SM-155M/DSN),
## 9 Samsung
## 10
## 11 OnePlus Nord N30 5G | Unlocked Dual-SIM Android Smart Phon
## 12
## 13 SAMSUNG Galaxy S21 FE 5G Cell Phone, Factory Unlocked Android Smartphon
## 14
## 15 TracFone | Motorola Moto g
## 16 Samsung Galaxy S22 Smartphone, Factory Unlocked Android Cell Phone, 128GB, 8K Camera
## 17 OnePlus Nord N200 | 5G Unlocked U.S Version | 6.49\" Full HD+LCD Screen | 90Hz Smo
## 18 OnePlus Nord N30 5G | Unlocked Dual-SIM Android Smart Phon
## 19 Google Pixe
## 20 Samsung Galaxy A15 (SM-155M/DSN),
## 21
```

## 22		TracFone   Motorola Moto g
## 23	Alcatel 1 (16GB) 5.0" Full View Display, Removable	
## 24		Samsung Galaxy A25
## 25		Motorola Moto G54 Dual-SIM 256GB RO
## 26		Motorola Moto
## 27		Google Pixel 8 - Unlocked
## 28		
## 29	OnePlus 9 Winter Mist, 5G Unlocked Android Smartphone U.S Version, 8GB RAM+128GB Storage, 120Hz F	
## 30		Google Pixel 9 - Unlocked Androi
## 31	Xiaomi Redmi Note 13 PRO 5G + 4G LTE (256GB + 8GB) 6.67" 200MP Triple (Tmobile M	
## 32	SAMSUNG Galaxy A25 5G A Series Cell Phone, 128GB Unlocked Android Smartphone, AMOL	