

# RWorksheet#5

Barrientos, Delfin, Infiesto

2024-11-06

```
library(polite)
library(rvest)
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)

url <- "https://www.imdb.com/chart/toptv/?ref_=nv_tvv_250"

session <- bow(url, user_agent = "Educational")

page <- scrape(session)

title_list <- page %>%
  html_nodes("h3.ipc-title__text") %>%
  html_text()

title_list
```

```
## [1] "IMDb Charts"
## [3] "2. Planet Earth II"
## [5] "4. Band of Brothers"
## [7] "6. The Wire"
## [9] "8. Blue Planet II"
## [11] "10. Cosmos: A Spacetime Odyssey"
## [13] "12. Our Planet"
## [15] "14. Bluey"
## [17] "16. Fullmetal Alchemist Brotherhood"
## [19] "18. Life"

"1. Breaking Bad"
"3. Planet Earth"
"5. Chernobyl"
"7. Avatar: The Last Airbender"
"9. The Sopranos"
"11. Cosmos"
"13. Game of Thrones"
"15. The World at War"
"17. Rick and Morty"
"19. The Last Dance"
```

```
## [21] "20. The Twilight Zone"          "21. The Vietnam War"
## [23] "22. Sherlock"                  "23. Attack on Titan"
## [25] "24. Batman: The Animated Series" "25. The Office"
## [27] "Recently viewed"
```

```
title_list_sub <- as.data.frame(title_list[2:251])
head(title_list_sub,50)
```

```
##           title_list[2:251]
## 1           1. Breaking Bad
## 2           2. Planet Earth II
## 3           3. Planet Earth
## 4           4. Band of Brothers
## 5           5. Chernobyl
## 6           6. The Wire
## 7           7. Avatar: The Last Airbender
## 8           8. Blue Planet II
## 9           9. The Sopranos
## 10          10. Cosmos: A Spacetime Odyssey
## 11          11. Cosmos
## 12          12. Our Planet
## 13          13. Game of Thrones
## 14          14. Bluey
## 15          15. The World at War
## 16          16. Fullmetal Alchemist Brotherhood
## 17          17. Rick and Morty
## 18          18. Life
## 19          19. The Last Dance
## 20          20. The Twilight Zone
## 21          21. The Vietnam War
## 22          22. Sherlock
## 23          23. Attack on Titan
## 24          24. Batman: The Animated Series
## 25          25. The Office
## 26          Recently viewed
## 27          <NA>
## 28          <NA>
## 29          <NA>
## 30          <NA>
## 31          <NA>
## 32          <NA>
## 33          <NA>
## 34          <NA>
## 35          <NA>
## 36          <NA>
## 37          <NA>
## 38          <NA>
## 39          <NA>
## 40          <NA>
## 41          <NA>
## 42          <NA>
## 43          <NA>
## 44          <NA>
## 45          <NA>
```

## 46	<NA>
## 47	<NA>
## 48	<NA>
## 49	<NA>
## 50	<NA>