

```
library(tidyverse)
library(rvest)
```

```
Warning message in system("timedatectl", intern = TRUE):
```

```
"running command 'timedatectl' had status 1"
```

```
Warning message:
```

```
"Failed to locate timezone database"
```

```
— Attaching packages — tidyverse 1.3.1
```

```
✓ ggplot2 3.3.5    ✓ purrr  0.3.4
✓ tibble  3.1.5    ✓ dplyr  1.0.7
✓ tidyr   1.1.4    ✓ stringr 1.4.0
✓ readr   2.0.2    ✓ forcats 0.5.1
```

```
— Conflicts — tidyverse_conflicts()
```

```
✗ dplyr::filter() masks stats::filter()
✗ purrr::flatten() masks jsonlite::flatten()
✗ dplyr::lag()     masks stats::lag()
```

```
Attaching package: 'rvest'
```

```
url <- "https://www.imdb.com/search/title/?groups=top_100&sort=user_rating,desc"
```

```
imdb <- read_html(url)
```

```
#movie title
titles <- imdb %>%
  html_nodes("h3.lister-item-header") %>%
  html_text2()
```

```
ratings <- imdb %>%
  html_nodes("div.ratings-imdb-rating") %>%
```

```
html_text2 %>%  
as.numeric()
```

```
ratings[1:10]
```

```
9.3·9.2·9·9·9·9·9·8.9·8.8·8.8
```

```
df <- data_frame(  
  title = titles,  
  rating = ratings  
)
```

```
head(df)
```

A tibble: 6 × 2

title	rating
<chr>	<dbl>
1. The Shawshank Redemption (1994)	9.3
2. The Godfather (1972)	9.2
3. The Dark Knight (2008)	9.0
4. The Lord of the Rings: The Return of the King (2003)	9.0
5. Schindler's List (1993)	9.0
6. The Godfather Part II (1974)	9.0

Warning message:

“`data_frame()` was deprecated in tibble 1.1.0.

Please use `tibble()` instead.

This warning is displayed once every 8 hours.

Call `lifecycle::last_lifecycle_warnings()` to see where this warning was gener

```
library(tidyverse)  
library(rvest)
```

Warning message in system("timedatectl", intern = TRUE):

“running command 'timedatectl' had status 1”

Warning message:

```
"Failed to locate timezone database"
```

```
— Attaching packages — tidyverse 1.3.1
```

```
✓ ggplot2 3.3.5    ✓ purrr  0.3.4  
✓ tibble  3.1.5    ✓ dplyr  1.0.7  
✓ tidyr   1.1.4    ✓ stringr 1.4.0  
✓ readr   2.0.2    ✓ forcats 0.5.1
```

```
— Conflicts — tidyverse_conflicts()
```

```
✗ dplyr::filter() masks stats::filter()  
✗ purrr::flatten() masks jsonlite::flatten()  
✗ dplyr::lag() masks stats::lag()
```

```
Attaching package: 'rvest'
```

```
url <- read_html("https://specphone.com/OPPO-A17k.html")
```

```
attributes <- url %>%  
  html_nodes("div.topic") %>%  
  html_text2  
  
values <- url %>%  
  html_nodes("div.detail") %>%  
  html_text2
```

```
df <- data_frame(  
  attribute = attributes,  
  value = values  
)  
df
```

A tibble: 31 × 2

attribute	value
<chr>	<chr>
วันเปิดตัว	พฤศจิกายน 2565
วันวางจำหน่าย	ยังไม่วางจำหน่าย
ขนาด	164.20 x 75.60 x 8.30 มม.
น้ำหนัก	189 กรัม
วัสดุ	ไม่รองรับ
SIM	รองรับ 2 ซิมการ์ด (nano sim, nano sim)
Technology	HSPA 42.2/5.76 Mbps, LTE
2G	850/900/1800/1900
3G	850/900/1900/2100
4G	850/900/1900/2100/2600
5G	-
ความเร็ว	HSPA 42.2/5.76 Mbps, LTE
ประเภท	IPS LCD
ขนาดหน้าจอ	6.56 นิ้ว
ความละเอียด	720 x 1612 pixels
ระบบปฏิบัติการ	Android 12
ชิปประมวลผล	MediaTek Helio G35 2.3 GHz
ชิปกราฟิก	PowerVR GE8320
หน่วยความจำ	3 GB
ความจุ	64 GB
Memory Card	ไม่รองรับ
กล้องหลัก	ตัวที่ 1: 8MP, f/2.0
ความละเอียดวิดีโอ	1080p@30fps
กล้องหน้า	ตัวที่ 1: 5 MP, f/2.2, (wide)
Bluetooth	5.3, A2DP
Wi-Fi	802.11 a/b/g/n/ac, dual-b
USB	micro USB
GPS	A-GPS
NFC	ไม่รองรับ
ความจุ	5,000 mAh
ประเภท	Non-removable Li-Po Batt

```
samsung_url <- read_html("https://specphone.com/brand/Samsung")
```

```
links <- samsung_url %>%  
  html_nodes("li.mobile-brand-item a") %>%  
  html_attr("href")
```

```
full_links <- paste0("https://specphone.com", links)
```

```
result <- data.frame()  
  
for (link in full_links) {  
  ss_topic <- link %>%  
    read_html() %>%  
    html_nodes("div.topic") %>%  
    html_text2()  
  
  ss_detail <- link %>%  
    read_html() %>%  
    html_nodes("div.detail") %>%  
    html_text2()  
  
  tmp <- data.frame(attribute = ss_topic, value = ss_detail)  
  
  result = bind_rows(result, tmp)  
}  
  
print(result)
```

```
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612
```

613

614

615

616

617

618

619

620

```
write_csv(result, "result_ss_phone.csv")
```