

RWorksheet_Lomibao#5(lomibao and rabago)

lomibao and rabago

2024-11-11

1.

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

url <- "https://www.imdb.com/chart/toptv/"

session <- bow(url,
               user_agent = "Student education purpose")
session

## <polite session> https://www.imdb.com/chart/toptv/
##   User-agent: Student education purpose
##   robots.txt: 35 rules are defined for 3 bots
##   Crawl delay: 5 sec
##   The path is scrapable for this user-agent
page <- read_html(url)

# titles <- page%>%html_nodes(".")%>%html_text()
#rating <- page%>%html_nodes(".")%>%html_text()
#voted <- page%>%html_nodes(".")%>%html_text()
#number_episode <- page%>%html_nodes(".")%>%html_text()
#year<- page%>%html_nodes(".")%>%html_text()
```