Mini Project 1 : IMDB web scraping

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

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

```
movies_name <- imdb %>%
    html_nodes("h3.lister-item-header") %>%
    html_text2()

ratings <- imdb %>%
    html_nodes("div.ratings-imdb-rating") %>%
    html_text2()

num_votes <- imdb %>%
    html_nodes("p.sort-num_votes-visible") %>%
    html_text2()
```

```
df_imdb <- data.frame(moveiname = movies_name,
            rating = ratings,
            vote = num_votes)
```

```
write_csv(df_imdb,"dataframeimdb.csv")
```

Mini Project2 : iphone model web scraping

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

```
url <- read_html("https://specphone.com/brand/Apple")
```

```
links <- 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) {
  models <- link %>%
    read_html() %>%
    html_nodes("h1.page-topic") %>%
    html_text2()

  topics <- link %>%
    read_html() %>%
    html_nodes("div.topic") %>%
    html_text2()

  details <- link %>%
    read_html() %>%
    html_nodes("div.detail") %>%
    html_text2()

  df_apple_model <- data.frame(model = models,
                               topic = topics,
                               detail = details)

  result <- bind_rows(result , df_apple_model)

}
```

```
write.csv(result,"applemodel.csv")
```