**PTT某個版**

* The link of the interested forum

url <- "https://www.ptt.cc/bbs/NTU/index2131.html"

* The xpath to crawl multiple news links  the first page of search result

linkpath <- "//\*[@id="main-container"]/div[2]//div/div[3]/a"

* The code to get back the title, content, journalist, and timestamp of each news

 author<-  trimws(xml\_text(xml\_find\_all(tdoc, '//\*[@id="main-content"]/div[1]')))

 title <- trimws(xml\_text(xml\_find\_first(tdoc, '//\*[@id="main-content"]/div[3]/span[2]')))

 content <- trimws(xml\_text(xml\_find\_all(tdoc, '//\*[@id="main-content"]')))

 timestamp <-  trimws(xml\_text(xml\_find\_all(tdoc, '//\*[@id="main-content"]/div[4]')))

* Screenshot of your dataframe (with at least 10 observations)

