

# R 語言 作業四重點回顧

王貿

國立臺灣大學行為與資料科學研究中心助理研究員

國立臺灣大學政治學系博士、兼任講師

maowang01@gmail.com

# str\_replace\_all 不能取代為NA

- data\_token %>%  
 mutate(  
 token = str\_replace\_all(  
 string = token,  
 pattern = "[0-9]+", replacement = NA\_character\_  
 )  
 )

# 停用詞data frame怎麼加？

- `tibble()`
- `data.frame()`：文字資料預設會變成factor
- `rbind`
- `dplyr::bind_rows`

## 小圖token依照次數排序

- `mutate(word_year = paste(word, year, sep =  
"_" ))`
- `ggplot(data,  
aes(x = reorder(token, word_n),  
y = word_n,  
fill = year))`

# theme 與 各種預設的theme\_xxxx

- 用預設的要注意，如果前面已經刪掉legend，在使用預設的theme，legend可能又會變回來。

# 移除圖標 ( legend )

- 有幾種做法都可以達到這個目的。
- `geom_bar(stat = "identity", show.legend = FALSE)`
- `geom_col(show.legend = FALSE)`
- `theme(legend.position = "none")`

# 請小心解讀資料

- 2016年發生什麼事？審計部為什麼會有方向上的調整？
- 請在提出資料無法佐證的意見時，補充其他資料支撐你的論點。

