

上課前你可以先.....

- 下載4個資料檔[trump_twitter_archive.csv, 兩個NTUSD開頭的檔案, mfd2.dic]與Rmd檔。
- 如果你還有想告訴我的意見，請填課後回饋調查（ <https://forms.gle/3aP9MPpEvsKGmbuh6> ）。

R 語言 情緒分析

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

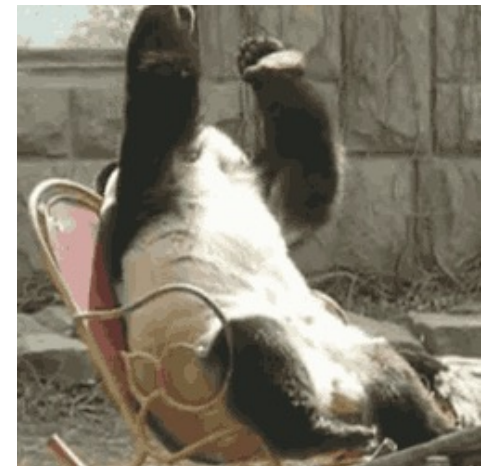
maowang01@gmail.com

課程主題重點

- 中英文情緒辭典 (`textdata`, `sentimentr`)
- 如何標註情緒到語料庫

辭典為基準（dictionary-based）的方法

- **文字是有情緒**的（但也有很多文字沒有情緒），不同的情緒辭典能提供不同的情緒分類與標註。
- 諸如停用詞，其實也是辭典為基準的應用。
- 只是先前是**以停用詞來移除觀察值**（anti_join），現在是以**情緒辭典加註情緒資訊**（inner_join）。



辭典為基準的缺點

- 無法處理多義詞。