

分手後我有話要說：

分析 PPT 男女版上近十年的文章內容



❖ G 0 2 匿名鴨子



➤ 林佳蓉

➤ 佟鈺麒

➤ 程依健

➤ 劉子安

目錄

一、前言

二、簡介

三、方法

1. 資料取得

2. 原始碼運作順序與說明

四、分析結果

1. 文字雲

3. 分手理由表

2. 交往時間長條圖

4. 分手情緒長條圖

五、結論

六、附錄

1. 組員分工

前言

在一起的時光如此長久甜蜜，是甚麼原因讓他們走到了分手的這個局面呢？

♡ 在一起越久越容易接納他的缺點，還是會越來越厭煩？
- - 交往多久越容易分手？七年？

♡ 相愛容易相處難。愛情探討眾說紛紜，但衝突與分手總是最大的風險

- - 究竟是什麼理由導致分手？個性不合？

♡ 分手後，彼此的分離會使人自由，還是會寂寞？
是樂觀，還是一蹶不振？

- - 分手後的情緒是什麼？負面的？

簡介

戀愛是當代社會生活中不可或缺的一部份。戀愛作為家庭的前奏曲而被賦予了神聖的想像，現實卻經常不盡人意，分手似乎便是這種理想與現實之間落差的證明。我們希望透過分手詞語分析，能夠得知在戀愛中那些地方容易出差錯，並且在這些地方能夠更加小心；也能建立自己對於分手的預期心理，讓我們能準備好面對分手的可能、以及可能會有的心情。

因此，我們提出了三個問題分別對應三個面向，想要來分析分手行為：**交往時間**、**分手理由**、**與分手情緒**。

彼此之間可能互有關連，但受限於我們的技術與所取得的情緒詞典內容的不完整，最終決定不進行交叉分析。

方法

1、資料取得

我們分析 **ptt 男女版** 標題有分手的發文內容，限定分類為**心情**，並**排除回文**。

ptt 男女版 俗稱分手版，因為發文內容通常都是戀愛中的衝突，且回文者通常都會鼓勵發文者分手。因此我們的分析對象可以說是具有特殊代表性。然而這些對象通常只包括異性戀，也因此有其研究侷限。選取心情且非回文，是為了確保選取到的文章發文者都是分手者，並且發的內容是有關於自己分手的情緒抒發。這能讓我們的情緒詞分析能更準確。

我們透過 **PTTmineR** 套件抓取帶有上述條件的文章，時間從 **2019/12/09** 到 **2007/07/31**，總共 **823** 篇文章。

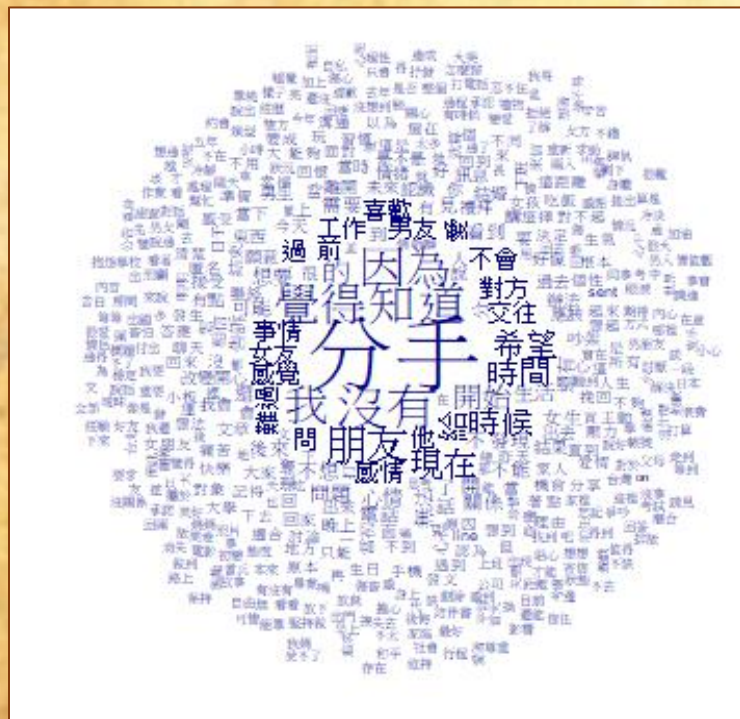
2、原始碼運作順序與說明

- 爬蟲
 - 根據資料條件來整理資料
 - 斷詞：三種斷詞，content 欄都包括了標題和內文的文字，也都以 jiabaR 來分類
 - ① 一般斷詞（位於 pttmine）
 - ② 排除掉只有一個字的詞的斷詞（pttmine）
 - ③ 詞性斷詞（pttminetag）
 - 建立情感詞統計：結合詞性斷詞與情感詞典
-
- 繪出文字雲：兩種文字雲
 - ① 以排除掉只有一個字的詞的斷詞繪圖
 - ② 以詞性為非介係詞的斷詞繪圖
 - 分析交往時間：統計「交往」詞後五個詞以內的時間單位並繪出長條圖
 - 分析分手理由：取得「因為」詞後十個詞以內的詞語，透過研究者主觀分析來找出主要理由
 - 分析分手情緒：以之前建立的情感詞統計來分析
 - ① 分手情緒詞的正負統計（emotion_dict）
 - ② 分手情緒詞的形容詞、名詞、動詞統計（emotion_vis）

[illegible]

1.1 排除掉一個字元的文字雲

連接詞與代名詞占大成，但可以看出我們的用量大於我



1.2 排除掉介系詞的文字雲

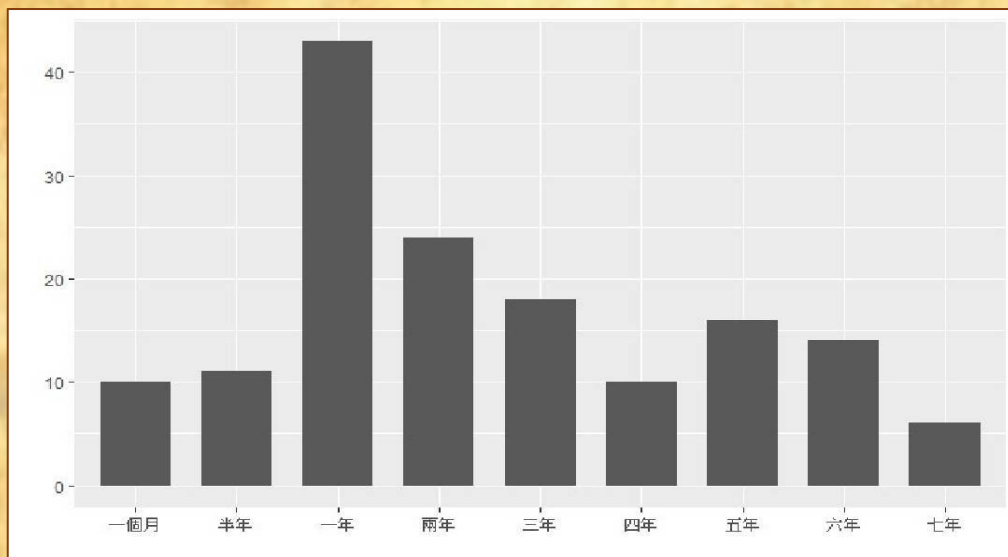
分類方法所以也把代名詞去掉。但能看出朋友、希望、時間等名、動詞增加

2.分手理由表(前十)

自己	93	自己提分手	朋友	45	分手後當朋友
覺得	82	覺得分手快樂	時間	36	分手過....的時間
沒有	76	沒有原因 沒有誰對誰錯	吵架	29	一直吵架→分手
知道	73	知道原因 知道不適合	家人	27	把她當家人 家人反對
工作	63	異地工作	遠距離	23	遠距離比較辛苦

自己、覺得、沒有、知道、朋友可能算不上典型的理由(特定事件)。值得一提的是，**個性不合**不在表內。

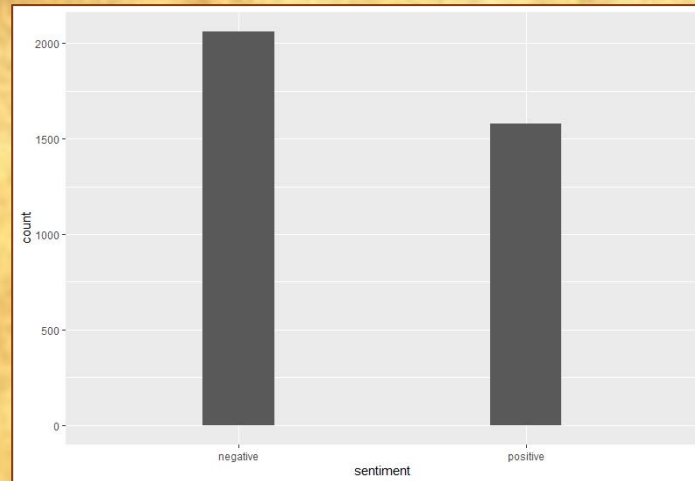
3.交往時間長條圖



以斜率來看，交往一年分手的人最多，其後的斜率一致。這代表著**交往一年的分手率最高**，其後的分手率則差不多，而交往一年內的分手率最低。

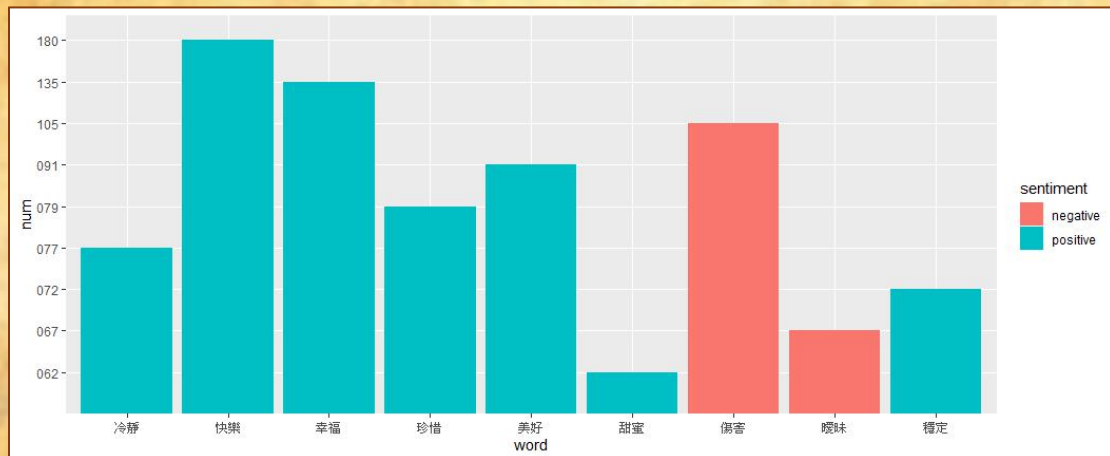
4.分手情緒長條圖

4.1 正負情緒詞



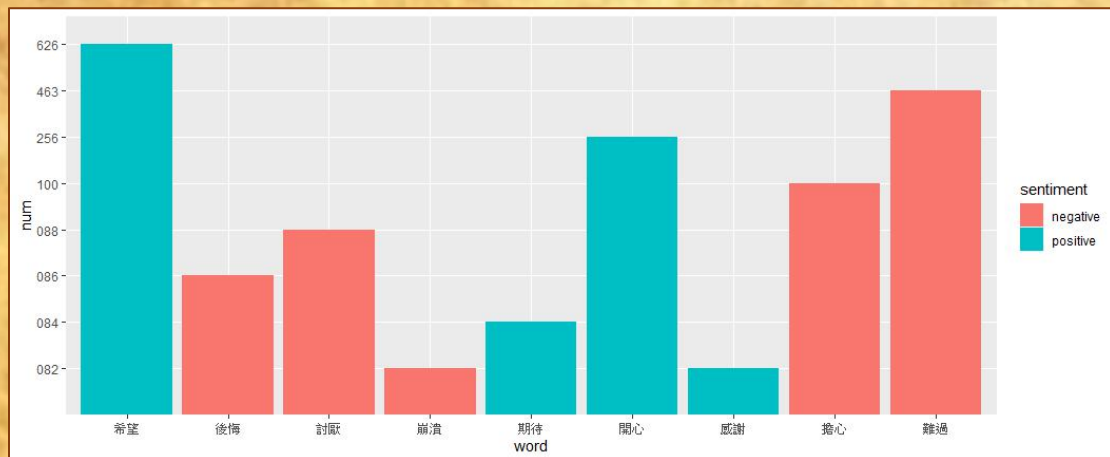
負面大於正面。不過這裡的情緒詞還包括代名詞，因此以這來詮釋分手情緒的意義並不準確。

4.2 形容詞情緒詞



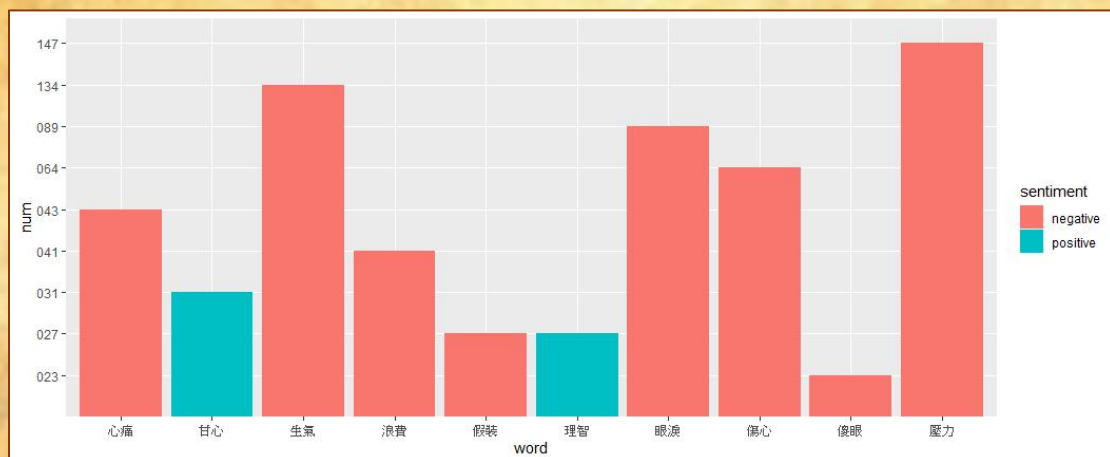
大多數為正面，「快樂」最多。其中回憶過去的詞占了快一半，包括幸福、珍惜、美好、甜蜜。其中冷靜與曖昧的情緒分類在分手的脈絡中可能是相反的。但我們在這裡仍保留情緒詞典原本的設計。

4.3 動詞情緒詞



負面較多。這裡的詞是我們原本預期可能會出現的分手後的心情。難過之於開心，崩潰之於希望。可能也因此比較適合去解釋我們對於分手的第三個問題：分手後的情緒是正面還是負面？

4.4 名詞情緒詞



大多數為負面，「壓力」最多。這裡的詞性是以 jiabaR 的 tag 功能來區分，但生氣、心痛、浪費、假裝、傷心、傻眼似乎更適合被列為動詞。如果將這部分納入動詞情緒詞統計，那麼動詞情緒詞便會以負面為大宗。

結論

Q 交往多久越容易分手？

A 一年或一年多

Q 究竟是什麼理由導致分手？

A 依序為異地工作，吵架，家人反對或家人感，遠距離。然而當事人也有很大的可能會找不到分手的原因。

Q 分手後的情緒是什麼？

A 過去的美好與現在的難過、壓力夾雜在一起，但希望、開心與快樂經常是勝過了這些。

附錄

1 · 組員分工

❖ 林佳蓉

- ① 情緒詞統計與正負情緒長條圖
emotion_dict.R

❖ 程依婕

- ① 爬蟲 data.Rmd
- ② 斷詞與文字雲
pttmine.Rmd & pttminetag.Rmd
- ③ 交往時間統計長條圖 find_love_time.Rmd
- ④ 書面報告

❖ 佟鈺麒

- ① 分手理由統計 find_reason.Rmd

❖ 劉子安

- ① 依詞性 情緒統計長條圖 emotion_vis.rmd
- ② 海報報告