|  |
| --- |
| 社群媒體分析  第一次讀書會作業\_第十組 |
| 一、分析主題：民眾對棒球經典賽國手張育成的愛恨情仇 |
| 二、組員：林育伈 N114320009 何軒育 N104220016  周嵩智 N104220009 許佳薇 N104220019  許耀文 N104220023 林素芬 N104220024  楊蕙瑄 N104220027 |
| 三、分析工具：中山大學工作流程平台 工作流程名稱：第十組\_HW1 |
| 四、動機：  中華隊國手張育成堪稱本屆世界棒球經典賽的靈魂人物。  在大聯盟旅美台將中更是改寫多項紀錄的最強打者，也是目前台灣在ＭLB唯一的現役選手。  當台灣球迷引頸期盼這場睽違五年的2023經典賽時，中華隊徵召張育成加入國家隊卻遭他婉謝邀請。但列管球員身份的他不應婉拒國家徵召，使得球迷不滿狂酸他「逃兵」。  在中華職棒會長蔡其昌親自出馬與張育成深談後，他才允諾加入中華隊。而經典賽開打期間，張育成用他的精彩表現打出令球迷一場場的感動比賽，也逆轉了球迷口中的「逃兵阿成」晉升為「國防部長」。 |
| 五、目的：  探討球迷對張育成的情緒正負面影響。 |
| 六、資料範圍，來源，關鍵字：  資料來源：蘋果日報sports(體育)  資料範圍：2022/10/01~2023/03/18  搜尋關鍵字：張育成  蘋果日報爬蟲任務結果： |
| 七、系統流程： |
| 八、分析過程  **替換字串**   * 文章內\n\n>>換成。 \n>>換成， * 任務結果：取代190筆   **中文斷詞設定**   * 參數設定:中華隊英雄名字“吉力吉撈”會斷詞為“吉力”及“吉撈”， “吳哲源” 會斷詞為“吳”及“哲源”   所以在以下詞增加權重經典賽 100、全壘打 100、吉力吉撈 100 、吳哲源 100  **清除停用詞**   * 參數設定：自定義停用字24個 * 任務結果：清除15550個停用字   **詞頻計算**  **Lexicon Based情緒分析**   * **參數設定**： 將”國防部長””三軍統帥”定義為正面詞彙。代表張育成經典賽贏球功勞之大。 將“逃兵”定義為負面詞彙。 * **任務結果**：一篇文章最大正向情緒17個,最大負向情緒有9個   **合併資料**   * 因要分別進行正負面情緒分析,所以需要利用system\_id欄位做為key值， 利用已分析出來的LexiconBased情緒分析與PTT原始文章來源進行join。   **欄位篩選**   * **參數設定**： 因後續分析是要挑出正負面情緒，所以只會用到sentiment\_Value及artDate兩個欄位 * **任務結果**   **資料篩選**   * **參數設定：**分析的是正面情緒,所以會將情緒分>0的篩選出來; 負面情緒,所以會將情緒分<0的篩選出來   **正(負)面情緒分析**   * **參數設定：**合併匯總group by 日期artDate,匯總函數用count去計算篇數(system\_id)   結果可看到有正面情緒的日期及篇數   * 任務結果：   **每日情緒總合**   * **參數設定：**合併匯總group by 日期artDate,匯總函數用sum，   原因是我們要看到每天的情緒值sentiment\_value的加總     * 任務結果：這裡可以看到每個日期下,那天所有文章的情緒加總分數   **LIWC情緒分析**      **MultiSentiClass**   * **參數設定：**增加欄位，此欄位抓出各個情緒數量欄位的最大值 * 任務結果：   **儀表板-折線圖**   * 利用折線圖可以看到球迷每天對張育成的愛恨情仇並對照出他們心情的喜怒哀樂(每日情緒總合) * 結論：2023/3/12以前情緒分數相對較低甚至出現負數，2023/3/12情緒分數最高。   **儀表板-文字雲**   * 利用文字雲來看詞頻計算的結果 * 張育成是經典賽台灣最熱門的的話題之一   **儀表板-直條圖** 利用直條圖來看正面情緒分析，在2023年3月12日的正面情緒最多  **儀表板-圓餅圖**  利用圓餅圖來看文章各個情緒佔比，發現文章以正面情緒最多。    **結論**：   1. 利用蘋果日報觀察民眾對於張育成的看法正面大於負面 2. 2023/3/12張育成正面情緒最高，也呼應當天中華隊對戰荷蘭隊打出致勝滿貫全壘打，掀起民眾對他的崇拜 3. 2023/3/12以前負面情緒相對較多，也呼應之前張育成原先想婉拒加入中華隊打球民眾憤怒的心情 |

**第十組謝謝老師&助教的指導！**

一張含有 人員, 布 的圖片

自動產生的描述