|  |
| --- |
| 房 市 觀 測 |
| 動機與結論：  疫情降至二級警戒後，隨著本土的2019冠狀病毒疾病（COVID-19）疫情趨穩，民眾恐慌情緒不再，遇上美國資金寬鬆，美元大幅貶值，資金流向商品市場，引發全球性通膨危機，台灣也多游離資金，資金潮環境持續，讓消費者預期房價上漲，央行雖再度出手管控房市，但不影響自用者，再加上近期國內、外通貨膨脹蠢蠢欲動，不動產資產將成為抗通膨的首選，故本研究將以房市做為本次分析議題，在研究的結果發現PTT房市版觀察民眾對於房市的看法是正面大於負面，另一件有趣的發現是10月5日一篇預售屋的詐騙案新聞，導致出現負面情緒分析的高峰。 |
| 資料範圍、來源、關鍵字：  資料來源：PTT home\_sale  資料範圍：2021/09/01~2021/10/21  搜尋關鍵字：房市  PTT爬蟲任務結果：總共有315篇與房市相關資料 |
| 系統流程： |
| 分析過程：  替換字串：  文章內有一些@@的符號，將它替換成空白    任務結果：取代了4筆    中文斷字設定  依據斷詞結果,反覆修正、清洗斷詞，將會用在房市的特殊用語做為斷詞，讓分析可以更為準確。  參數設定：    任務結果    清除停用詞：  由任務結果及詞頻計算結果去觀察文字，同樣反覆清洗及加入停用字。  參數設定：利用自定義停止詞進行去除雜訊文字    任務結果  觀察內容是否較符合所要資訊後，再進行下一步驟    詞頻計算  篩選詞彙數量設定為100    任務結果      Lexicon Based情緒分析  參數設定笑死及XD是為嘲笑的意思，故將其定義在負面詞彙    任務結果  一篇文章最大正向情緒42個,最大負向情緒有28個    合併資料  因要進行負面情緒分析,所以需要利用system\_id欄位做為key值，利用已分析出來的Lexicon Based情緒分析與PTT原始文章來源進行join。    欄位篩選  因後續分析是要挑出負面情緒，所以只會用到sentiment\_Value及artDate兩個欄位    篩選結果  以下是每篇文章中情緒的分數，情緒值小於0即為負面情緒。    資料篩選  我們要分析的是負面情緒,所以會將情緒分數<0的篩選出來。    篩選結果    負面情緒分析  合併匯總group by 日期artDate,匯總函數用count去計算篇數(system\_id)  結果可看到有負面情緒的日期及篇數    負面情緒分析-結果    每日情緒總合  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    儀表板-折線圖  利用折線圖可以看到鄉民每天對房市的愛恨情仇並對照出他們心情的喜怒哀樂(每日情緒總合)    儀表板-文字雲  利用文字雲來看詞頻計算的結果  房價、房市的火熱可以由此看出，地點為高雄、台南、台中，也可以看到政府跟央行有所行動(打房)，所以也出現在文字雲中。    儀表板-直條圖  利用直條圖來看負面情緒分析，在10月4日及5日的負面情緒最多    到10/5探究原因,發現system\_id:214有11個負面用詞,206有6個    新增「查看異常資料流程」  再自訂加入一個查看異常資料流程，將system\_id 設定為214  原來是幽靈預售屋的新聞，大輪建設總經理用此波熱潮詐騙投資人，收了投資人4戶約近4千萬元頭期款，預售屋詐騙脫產空殼，導致負面能量滿滿。      儀表板-直條圖  利用直條圖來看正面情緒分析，發現928檔期史上最滿(928檔期 是預判下半年房市溫度指標，也是延續到明年的整體房市動向參考依據。)  結論：  利用PTT房市版觀察民眾對於房市的看法是正面大於負面  結論：疫情趨緩、美國資金寬鬆，美元大幅貶值，資金流向商品市場，引發全球性通膨危機，台灣也多游離資金，皆為房市回溫原因。    第六組成員：  N094020022 李昀瑾  N094220011 吳家羚  N094220016 林映汝  N094220021 李宛珊  N094220024 吳芷潁  N094220025 蘇則融  N094220029 王莉筑 |