1.請下載一篇近期之財經新聞，設定適當之停用詞，利用TF-IDF及TextRank兩模式進行關鍵詞分析，分析出前十個關鍵詞及其權重，並說明分析結果。

2.承上題，僅取名詞或動詞為關鍵詞，利用TF-IDF及TextRank兩模式進行關鍵詞分析，分析出前十個關鍵詞及其權重，並說明分析結果。

在做jieba段詞時有發現一些詞段的不是很理想所以增加以下停用詞。

增加的停用詞 : 分析師、美股、了結、雙十、國泰證期、聯準會、位階、個股、乖離、觸及。

從下兩張表可看出不管使用TF-IDF和TextRank的模式’反彈’和’趨勢’兩個關鍵詞都名列前茅。

在僅取名詞或動詞為關鍵詞可以看到’反彈’、’趨勢’、’破底’、’反轉’位居前四，搭配前面沒有限制詞性的關鍵詞可以知道在股市方面會有明顯的起伏，若要進行投資的投資人需要謹慎操作。

TF-IDF 僅取名詞或動詞為關鍵詞 allowPOS=('n','v')

| 關鍵詞 | 權重 | 關鍵詞 | 權重 |
| --- | --- | --- | --- |
| 反彈 | 0.340 | 反彈 | 0.621 |
| 美股 | 0.298 | 趨勢 | 0.466 |
| 趨勢 | 0.255 | 破底 | 0.388 |
| 蔡明翰 | 0.255 | 反轉 | 0.311 |
| 破底 | 0.213 | 數據 | 0.311 |
| 大漲 | 0.170 | 建議 | 0.311 |
| 反轉 | 0.170 | 指數 | 0.311 |
| 建議 | 0.170 | 就業 | 0.233 |
| 投資人 | 0.170 | 走勢 | 0.233 |
| 美國 | 0.128 | 認為 | 0.233 |

TextRank 僅取名詞或動詞為關鍵詞 allowPOS=('n','v')

| 關鍵詞 | 權重 | 關鍵詞 | 權重 |
| --- | --- | --- | --- |
| 趨勢 | 1.000 | 趨勢 | 1.000 |
| 反轉 | 0.740 | 反彈 | 0.705 |
| 反彈 | 0.712 | 反轉 | 0.664 |
| 指出 | 0.669 | 指出 | 0.661 |
| 破底 | 0.657 | 破底 | 0.647 |
| 認為 | 0.635 | 走勢 | 0.591 |
| 走勢 | 0.589 | 下跌 | 0.583 |
| 下跌 | 0.582 | 認為 | 0.582 |
| 風險 | 0.552 | 就業 | 0.493 |
| 投資 | 0.510 | 風險 | 0.466 |