這次的作業跟上個作業用了同個資料集，只是把模型從VSM換成BM25，BM25除了一樣考慮了tf和idf之外，還再加入了一個document length的要素，並且有更多的參數可以調整，也跑比較久。

調參的部分我直接搜尋常用的範圍，K1通常會落在1.2~2之間，而b就和講義上寫的一樣公認設0.75，不過講義上的公式和網上查到常用軟體Elasticsearch上的BM25公式長得不太一樣，所以我直接把K3設成跟K1一樣，不過最後做出來的分數並沒有比VSM高就是了。

雖然兩個模型的公式看起來很不一樣，但是核心概念感覺差別沒那麼大，主要仍是衡量「一個index term在一個document中的權重」、「一個index term在一個query中的權重」以及「該index term本身的鑑別度(資訊量)」，最後再用這三個元素衡量document跟query之間的關聯度然後排序。