

友 李賀 痘
邊塞 自然 浪漫 州 奇險 江南
岑參 閹珊瑚 酒 詠物 淚 元稹 明日
李賀 自然 浪漫 州 奇險 江南
閹珊瑚 酒 詠物 淚 元稹 明日
王維 雨 懷人 豪放 李白

數據唐詩

王姿嵐 陳騏業 林奕均

孟浩然 故國 社會 惆悵 悵望 杜牧 馬
愁 白居易 長安 思鄉 曹植 抒情 深處
旅 高適 王之渙 笑 韋應物 登高



前言與動機

詩詞曲向來被視為文學上的瑰寶；
本次專題透過資料分析找出【唐詩】中：
【詞彙之間】、【詞彙與詩人之間】的相關性
以探討詞彙之間的關聯程度，
及作者生平、性格與常用字詞的關係。

資料來源

• 詩詞名句網 <http://www.shicimingju.com/>

唐诗三百首(306首)

查询到306首诗词

■ 新窗口打开

全部 唐(306)

全部 李白(33) 韩翃(3) 岑参(7) 张九龄(5) 王维(30) 权德舆(1) 白居易(6) 祖咏(2)
杜甫(35) 李商隐(22) 马戴(2) 李颀(7) 王建(1) 王翰(1) 刘方平(2) 张籍(1) 杜牧(10)
陈子昂(1) 高适(2) 许浑(2) 王勃(1) 杜秋娘(1) 张继(1) 温庭筠(4) 韩愈(4) 孟浩然(14)
韩偓(1) 孟郊(2) 贺知章(1) 薛母潜(1) 元结(2) 唐玄宗(1) 杜审言(1) 刘禹锡(4) 王湾(1)
钱起(4) 李益(3) 司空曙(3) 张乔(1) 崔涂(2) 柳宗元(5) 韦庄(2) 王之涣(2) 僧皎然(1)
崔颢(3) 骆宾王(2) 崔曙(1) 皇甫冉(1) 刘慎虚(1) 卢纶(3) 薛逢(1) 王昌龄(9) 秦韬玉(1)
刘长卿(11) 常建(2) 韦应物(12) 李端(1) 张祜(4) 贾岛(1) 西鄙人(1) 邱为(1) 张旭(1)
柳中庸(1) 朱庆余(1) 戴叔伦(1) 郑畋(1) 陈陶(1) 宋之间(2) 益嘉运(1) 杜荀鹤(1) 元稹(2)
朱庆馀(1) 沈佺期(2) 张泌(1)

探討面向

資料視覺化

資料分析

【文字雲】 【長條圖】 【PCA/K-means】

- 詞彙之間 • 詞彙與詩人 • 詩人之間
- #TAG與詩人 • #TAG之間

第 1 款

- 詩人 / 詞彙

詩人與其代表詞彙

- 各詩人最常用的前十五大詞彙

第 2 款

- #主題字 / 詞彙

詞彙之間的鏈結

- 一主題詞彙容易帶動的其他詞彙

第 3 款

- #主題字 / 詩人

同一框架下的詩人

- 善於使用相同主題文字的詩人們

第 4 款

- 兩主題字 / 詞彙

內容相似的主題

第 5 款

- 兩詩人 / 詞彙

詩人與詩人之間

- 擁有相似詞彙的兩主題
- 擁有相似詞彙的兩詩人

STEP 1

資料集



唐詩文本蒐集

· 唐詩三百首

322 lines (321 sloc) 99.1 KB
1 name author content tag style
2 行宮 唐代：元稹 寥落古行宮，宮花寂寞紅。白頭宮女在，閒坐說玄宗。
3 登鸕雀樓 唐代：王之渙 白日依山盡，黃河入海流。欲窮千里目，更上一層樓。
4 新嫁娘詞 唐代：王建 三日入廚下，洗手作羹湯。未諳
5 相思 唐代：王維 紅豆生南國，春來發幾枝。願君多采擷，此物最相思。
6 雜詩三首·其二 唐代：王維 君自故鄉來，應知故鄉事。
7 鹿柴 唐代：王維 空山不見人，但聞人語響。返景入深林，
8 竹裡館 唐代：王維 獨坐幽篁裡，彈琴復長嘯。深林人不知，明月來相照。
9 送別/山中送別 唐代：王維 山中相送罷，日暮掩柴扉。
10 問劉十九 唐代：白居易 绿蟻新醅酒，紅泥小火爐。晚
11 哥舒歌 唐代：西鄙人 北斗七星高，哥舒夜帶刀。至今
12 靜夜思 唐代：李白 床前明月光，疑是地上霜。舉頭望
13 怨情 唐代：李白 美人卷珠簾，深坐暎蛾眉。但見淚痕
14 樂遊原/登樂遊原 唐代：李商隱 向晚意不適，驅車登古
15 聽箏 唐代：李端 鳴箏金粟柱，素手玉房前。欲得周郎
16 渡漢江 唐代：宋之間 嶺外音書斷，經冬復歷春。近鄉
17 八陣圖 唐代：杜甫 功蓋三分國，名成八陣圖。（名成
18 宿建德江 唐代：孟浩然 移舟泊煙渚，日暮客愁新。野
19 春曉 唐代：孟浩然 春眠不覺曉，處處聞啼鳥。夜來風
20 春怨/伊州歌 唐代：金昌緒 打起黃鸝兒，莫教枝上啼。
21 江雪 唐代：柳宗元 千山鳥飛絕，萬徑人蹤滅。孤舟蓑
22 秋夜寄邱員外/秋夜寄丘二十二員外 唐代：韋應物 懷君
23 終南望餘雪 唐代：祖詠 終南陰嶺秀，積雪浮雲端。林
24 宮詞·故國三千里 唐代：張祜 故國三千里，深宮二十年

· 詩人集

孟浩然.txt
孟郊.txt
張繼.txt
李世民.txt
李商隱.txt
李煜.txt
李白.txt
李賀.txt
杜牧.txt
杜甫.txt
柳宗元.txt
溫庭筠.txt
王勃.txt
王昌齡.txt



轉換類別型態、列出

· 整體分類（篇名/詩人/內文）

name	author	content
行宮	元稹	寥落古行宮，宮花寂寞紅。白頭宮女在，閒坐說玄宗。
登鸕雀樓	王之渙	白日依山盡，黃河入海流。欲窮千里目，更上一層樓。
新嫁娘詞	王建	三日入廚下，洗手作羹湯。未諳
相思	王維	红豆生南國，春來發幾枝。願君多采擷，此物最相思。
雜詩三首·其二	王維	君自故鄉來，應知故鄉事。來日綺窗前，寒梅著花未？
鹿柴	王維	空山不見人，但聞人語響。返景入深林，復照青苔上。
竹裡館	王維	獨坐幽篁裡，彈琴復長嘯。深林人不知，明月來相照。
送別/山中送別	王維	山中相送罷，日暮掩柴扉。春草明年綠，王孫歸不歸？
問劉十九	白居易	綠蟻新醅酒，紅泥小火爐。晚來天欲雪，能飲一杯無？
哥舒歌	西鄙人	北斗七星高，哥舒夜帶刀。至今窺牧馬，不敢過臨洮。
靜夜思	李白	床前明月光，疑是地上霜。舉頭望明月，低頭思故鄉。
怨情	李白	美人卷珠簾，深坐暎蛾眉。但見淚痕溼，不知心恨誰。
樂遊原/登樂遊原	李商隱	向晚意不適，驅車登古原。夕陽無限好，只是近黃昏。
聽箏	李端	鳴箏金粟柱，素手玉房前。欲得周郎顧，時時誤拂弦。
渡漢江	宋之間	嶺外音書斷，經冬復歷春。近鄉情更怯，不敢問來人。
八陣圖	杜甫	功蓋三分國，名成八陣圖。江流石不轉，遺恨失吞吳。
宿建德江	子浩然	移舟泊煙渚，日暮客愁新。野

STEP 2

程式內容

- ✓ 清理檔案
- ✓ 切詞，刪去符號、英文
- ✓ 建立TDM文本矩陣，統計辭彙出現數量
- ✓ 由TDM轉TFIDF，檢視圖表
- ✓ PCA + K-means 尋找相關性

STEP 3

發佈至 SHINY



An error has occurred

The application failed to start (exited with code 1).

```
Error in value[[3L]](cond) : there is no package called 'ggplot2'  
Calls: local ... tryCatch -> tryCatchList -> tryCatchOne -> <Anonymous>  
停止執行
```

～延伸～

- 輸入文字以尋找與你最相似的詩人

