知乎

指導教授

謝舒凱 教授

Team10 學生

B02611013 生機四 黃薇甄
B02704033 國企四 張景淵
B02705002 資管四 周晅瑢
B02705006 資管四 高偉立
B02705027 資管四 陳信豪
B02705028 資管四 冷俊瑩

INTRODUCTION

知乎

https://www.zhihu.com/

中國（世界）最大的 QA 問答社群

知乎 搜索你感兴趣的内容... 首页 话题 发现 消息 提问 陳豪

提问 回答 写文章 草稿

最新动态 设置

来自话题: 生活常识

经常吃速冻水饺好不好?

839 丁香医生，身体上的问题，来问丁香医生

在冬至的中午答这题，是比较应景的。 毕竟冬至是继元旦、除夕、春节、清明、夏至、端午、中秋、重阳.....之后，又一个吃饺子的节日。 质量合格的速冻饺子和现包的饺子最大的区别其实只有一点： 为了便于冷藏并保持比较好的口感，**速冻水饺往往添加了大量的油脂...** 显示全部

+ 关注问题 350 条评论 • 禁止转载

来自话题: 平面设计

有什么好看的银行卡?

756 招商银行信用卡客服中心，信用卡咨询/服务/乐活

感谢 @Fenng 邀请，其实一直在关注这个问题，终于我们能来填一填这个坑了，好饭不怕晚。 这个题目下已经有很多关于信用卡卡片的回答，但是有很多都是国外银行的信用卡，所以我们补充一下国内的招商银行信用卡。 声明：**此回答下的所有卡片图片版权归招商银行...** 显示全部

+ 关注问题 345 条评论 • 禁止转载

我的收藏

我关注的问题

邀请我回答的问题 283

公共编辑动态

社区服务中心

版权服务中心

知乎专栏

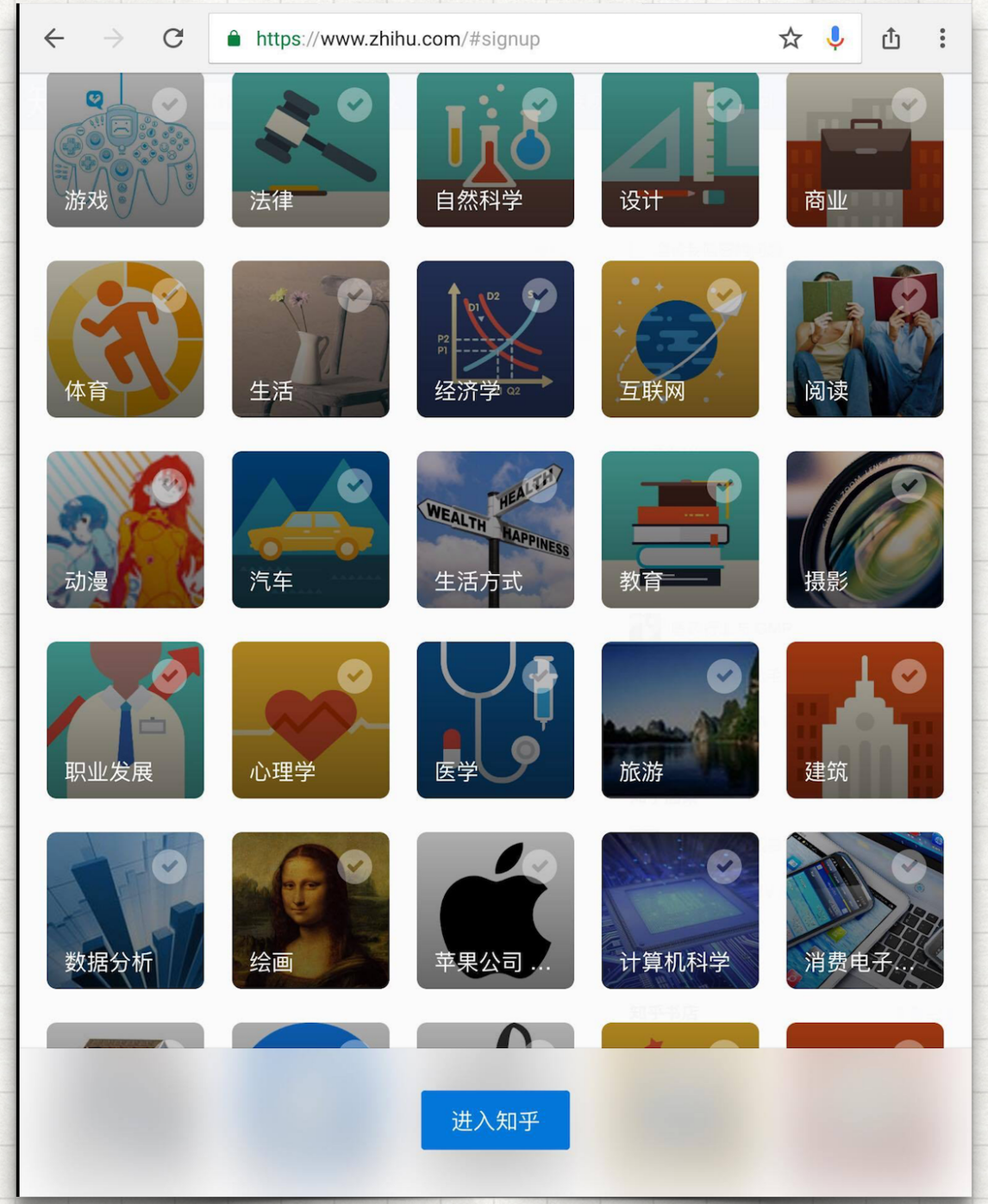
专栏 · 发现

知乎 Live 查看全部 »

中國的...

YAHOO! 奇摩 知識+

話題多元


https://www.zhihu.com/#signup
游戏
法律
自然科学
设计
商业
体育
生活
经济学
互联网
阅读
动漫
汽车
生活方式
教育
摄影
职业发展
心理学
医学
旅游
建筑
数据分析
绘画
苹果公司 ...
计算机科学
消费电子...
进入知乎

用戶眾多

千萬以上的使用戶

文本豐富

ONE ANSWER

知乎 搜索你感兴趣的内容... 提问

首页 话题 发现 消息 4 我

思想 社会 卡尔·马克思（Karl Marx） 社会主义女性主义 共产主义运动 修改

共产主义最喜欢反思，但为何反倒成了当今世界最弱势的意识形态？ 修改

所有人都说喜欢善于反思的人，共产主义最善于反思，可为啥连某些属于中世纪的意识形态都干不过呢？先是在伊朗被窃取革命果实，然后中亚和新疆也渐渐沦陷。

共产主义出极端者了，最先跳出来的就是温和派共产主义者，我们都知道黑共产主义最火的两本书《1984》和《动物庄园》是谁写的。我也同样也知道当年是谁灭了红色高棉。共产主义者内部的批评和自我批评几乎是所有意识形态中最激烈的，可为什么偏偏现在却处于最弱势的地位了呢？ 修改

31 条评论 分享 • 邀请回答 举报

135 个回答

中華電信 22:04

ofever wait， 主张改良解决我国问题的共产主义者

107 人赞同

笔者先理一下题主逻辑的反面：题主的大前提是-善于反思的意识形态强势。小前提是-共产主义意识形态很善于反思。结论是-共产主义是种强势意识形态。可是，某种思想善于反思就一定很强势吗？反思与某种思想意识形态是否强势，个人看来两者并没有多大的关系。倒是发现，很多强势的意识形态，例如民族主义，民粹主义，个人利己主义，恰恰是用不着怎么反思的。**也就是说，题主问了个风马牛不相及的坏问题。**

笔者认为，题主想问的是共产主义意识形态为何在现在会衰落。这个问题如果详细答的话足够写一整套丛书了。笔者尝试做一个很粗略的简述。

共产主义意识形态最辉煌的时期，是19世纪后期至20世纪70年代末。19世纪80、90年代，德国社会民主党在普选中去的辉煌成果，欧洲各国纷纷成立了社会主义的工人政党，晚年的恩格斯表示科学社会主义已经基本驳倒了其他各色社会主义流派。此后，俄国十月革命的胜利，苏联工业化和反法西斯战争的成就，东欧、亚洲、拉美社会主义国家的建立，民族解放运动的高潮以及其同社会主义阵营的亲和性质（为国家名称或宪法或党纲中加上“社会主义”一度成为流行），资本主义阵营对共产主义公开宣战。在那个时代，提“社会主义”就和现在提“自由民主人权宪政”一样，是进步、正义的象征。20世纪中叶，西方陷入石油危机的泥沼，而苏联则迎来了“停滞的黄金年代”与实力顶峰，中国国内的大革命与西方新社会运动以及反抗精神遥相呼应。

为什么能造成这样的局面？笔者认为大致有如下原因。

第一，马克思、恩格斯开创，列宁所继承的理论，确实揭露了当时资

资本主义好在哪？资本主义国家最终赢得了太空竞赛，在发展生产力方面抢先一步；资本主义国家通过其续命医生-社会民主党，在资本主

ONE QUESTION

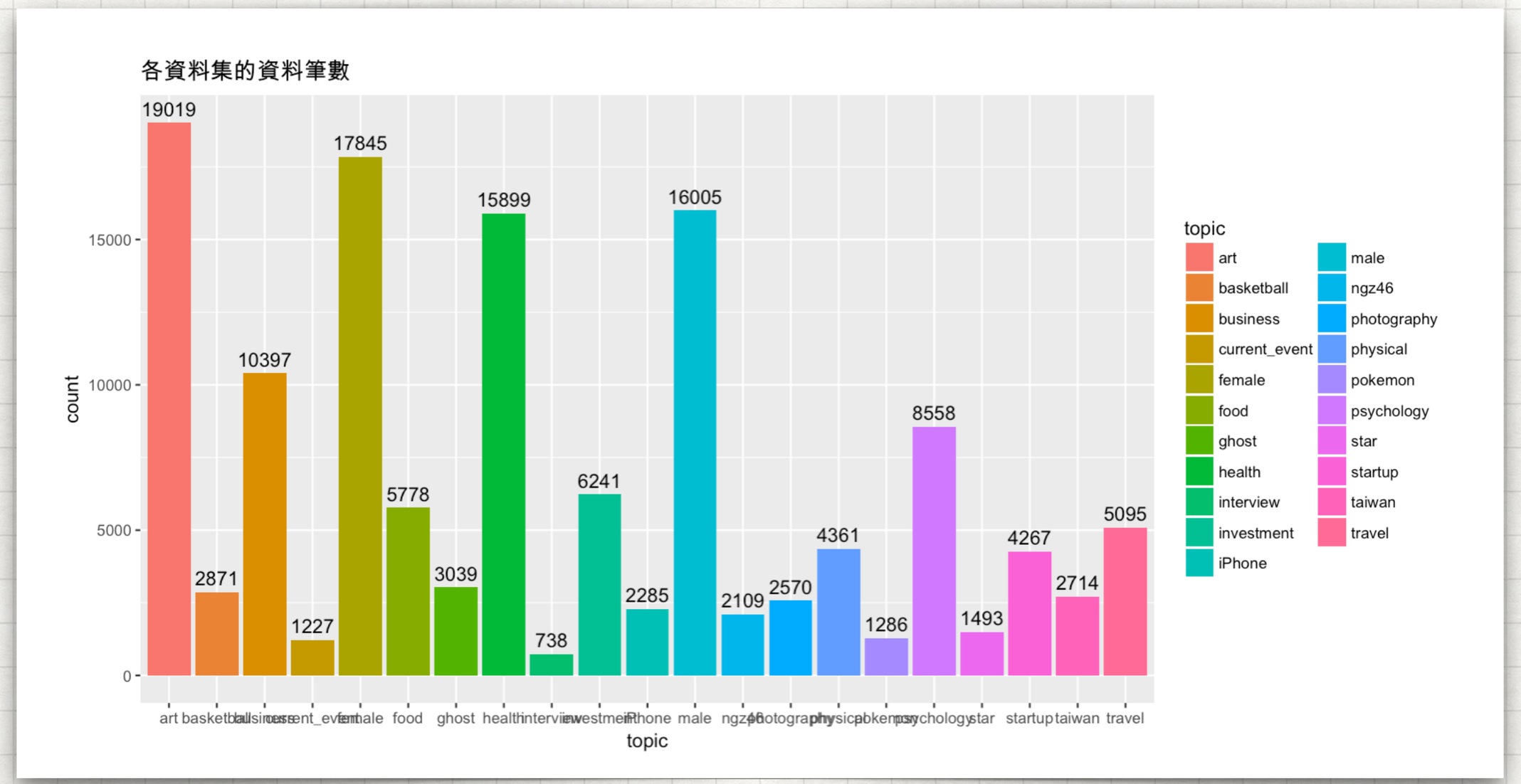
DATA

非知乎官方的 Python API

https://github.com/7sDream

選擇不同話題下的熱門問題來爬取資料

資料	說明
question_title	問題標題
question_detail	問題內容
question_time	發表問題時間
question_follower_num	問題追蹤人數
ans	回答內容
ans_time	發表回答時間
ans_upvote_num	回答贊同數
ans_collect_num	回答蒐藏數
ans_comment_num	回答評論數
author_follower_num	回答者被追蹤的人數
author_followee_num	回答者追蹤的人數
author_upvote_num	回答者獲得總讚數
author_thank_num	回答者獲得總感謝數
author_answer_num	回答者回答數
author_question_num	回答者發問數
author_post_num	回答者文章發表數
author_name	回答者名稱



21 份 DATASET

1,300,000 筆 DATA

17 種 VARIABLE

PREPROCESSING

文本前處理

請輸入要清理的文本

(個人調節）②參加附近行業的沙龍和論壇培訓等， 跟他人接觸靠別人來提升自己。 （圈子調節）③如果你有團隊那就團隊之間互相討論開會什麼的 （我就有個合伙人，是我高中好基友，連哄帶騙好歹給拉過來了）相互扶持相互努力，這樣快樂X2 痛苦÷2 真的非常非常感謝有朋友的一路相伴！ （團隊調節）<img src="https://pic4.zhimg.com/50/919cc17e9cf57882b9d47b73a8a9e2a7_b.jpg" data-rawwidth="357" data-rawheight="220" class="content_image" width="357">

濕貨說完了 該說重頭戲了 ：<u>乾貨</u>
技巧和方法 思路（這玩意我自己都沒啥）
這個是一個系統的科學，從頭開始。。。
 打字太多了 好累啊 休息休息 看看有沒有人跟著看吧…… <img src="https://pic3.zhimg.com/50/72a2133a1804c997387e4e9efd920cfa_b.jpg" data-rawwidth="96" data-rawheight="150" class="content_image" width="96">

 ++++++

clean_text()

過濾

網頁標籤、url、

過多的連續標點符號與空白

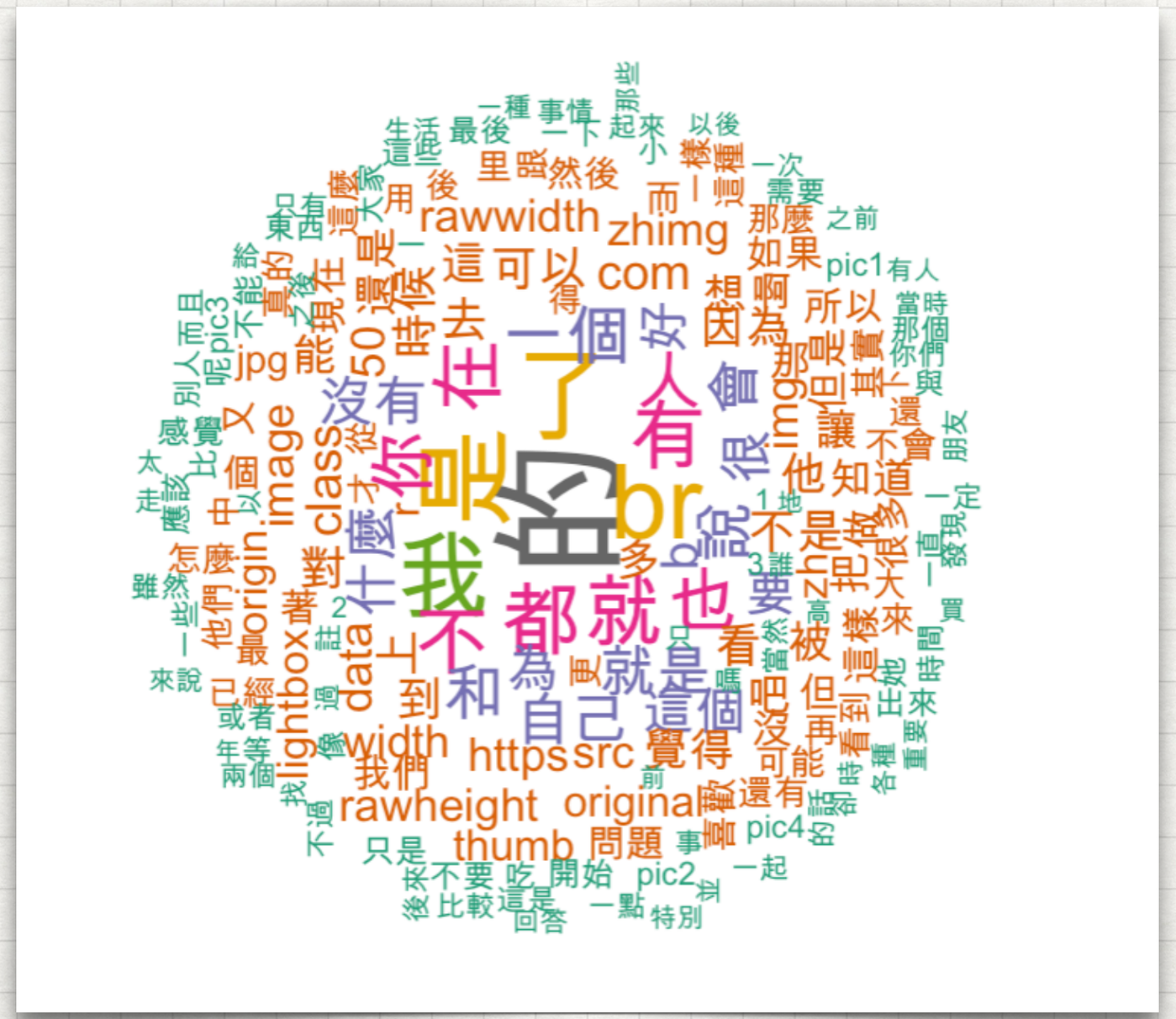
(個人調節）②參加附近行業的沙龍和論壇培訓等， 跟他人接觸靠別人來提升自己。 （圈子調節）③如果你有團隊那就團隊之間互相討論開會什麼的 （我就有個合伙人，是我高中好基友，連哄帶騙好歹給拉過來了） 相互扶持相互努力，這樣快樂X2 痛苦÷2 真的非常非常感謝有朋友的一路相伴！ （團隊調節）濕貨說完了 該說重頭戲了 ：乾貨技巧和方法 思路（這玩意我自己都沒啥） 這個是一個系統的科學，從頭開始 打字太多了 好累啊 休息休息 看看有沒有人跟著看吧 貌似被群嘲了 來吧 讓濃白的雨點來的更猛烈點

Jieba 斷詞後濾掉 stop word 的結果

調節 / 參加 / 附近 / 行業 / 沙龍 / 論壇 / 培訓 / 接觸 / 提升 / 圈子 / 調節 / 團隊 / 團隊 / 互相 / 討論 / 開會 / 合伙人 / 基友 / 連哄帶 / 騙 / 好歹 / 拉過來 / 相互 / 扶持 / 相互 / 快樂 / X2 / 痛苦 / 非常感謝 / 一路 / 相伴 / 團隊 / 調節 / 濕 / 貨 / 說完 / 重頭戲 / 乾貨 / 技巧 / 思路 / 玩意 / 沒啥 / 系統 / 科學 / 從頭開始 / 打字 / 太多 / 好累 / 休息 / 休息 / 看吧 / 貌似 / 群嘲 / 濃白 / 雨點 / 猛烈

STOP WORDS

取前 df 值最高的 2.5% 詞彙



TEXT TOPIC

OBSERVING

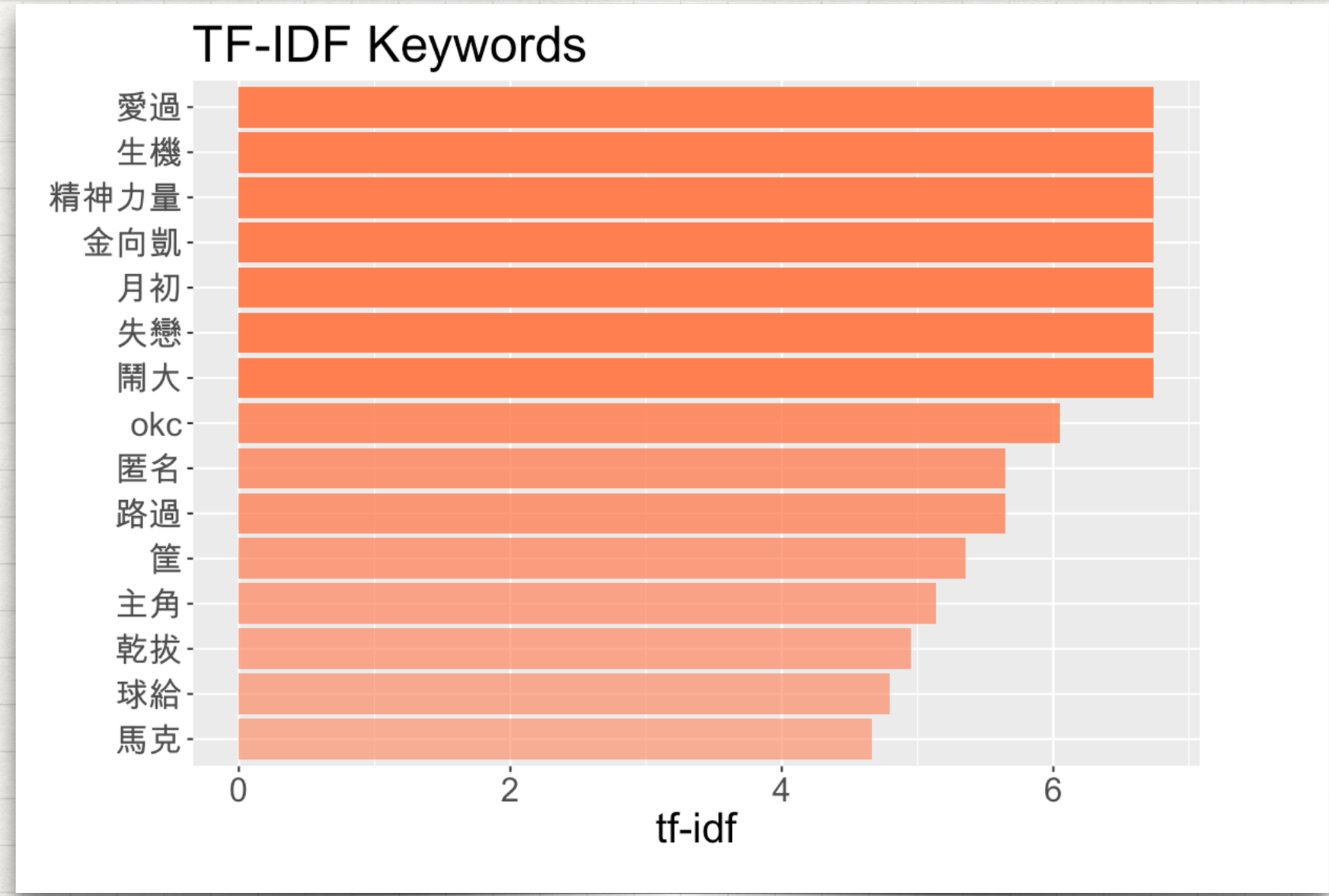
Word Cloud

請選擇一個主題
basketball
更新頁面
主題文字雲

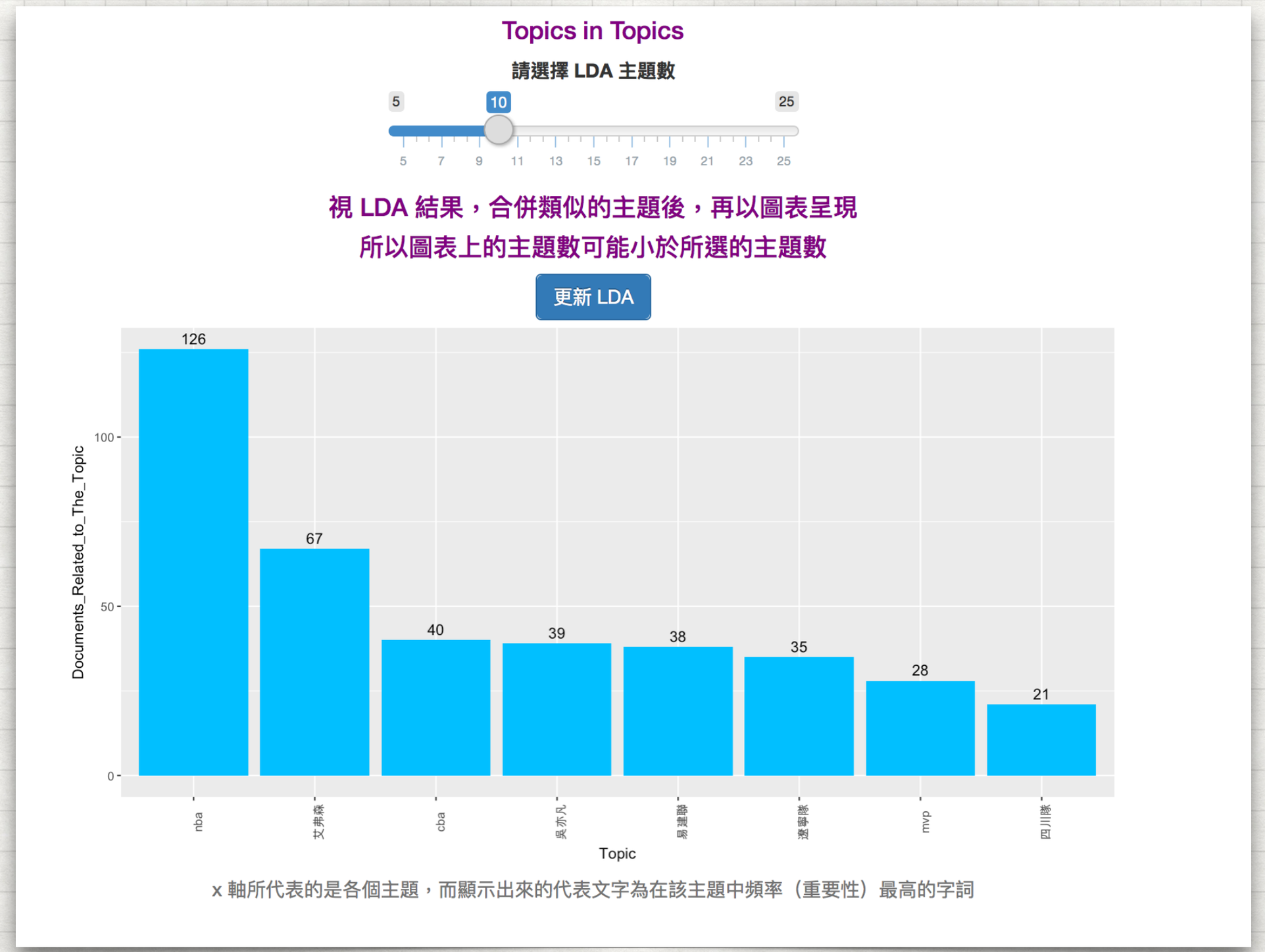
請選擇一個主題
star
更新頁面
主題文字雲

請選擇一個主題
health
更新頁面
主題文字雲

TF-IDF
Keywords



LDA
(Topics in Topics)



GOAL

知乎

既然你誠心誠意的發問了，我們就大發慈悲的回答你！

我們的專案

既然你誠心誠意的回答了，我們就大發慈悲的幫你打分數！

定義一篇回答文章的品質

(回應時間 = 回答發表時間 - 問題發表時間)

定義一篇回答文章的分數

你的回答在 SVM 模型裡
有多少比例接近 good Answer

SVM PROBABILITY (ONE AGAINST ALL)

Score: 73!

73%

good Answer

27%

bad Answer

CLUSTERING FEATURES

ESSAY SIMILARITY

SENTIMENT SCORE

SVM PROBABILITY (OAO)

STOP WORD COUNT

AUTHOR INFORMATION

WORD COUNT

CHARACTER COUNT

SENTIMENT

情感分數

你正在想什麼呢？

現在在台上的我，
既緊張又害怕又高興又興奮又不知所錯
我好像在胡言亂語，嗚嗚嗚

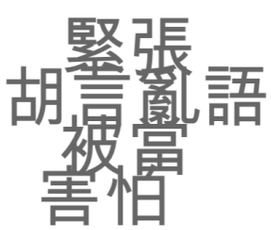
希望我不會被當QQ

告訴我

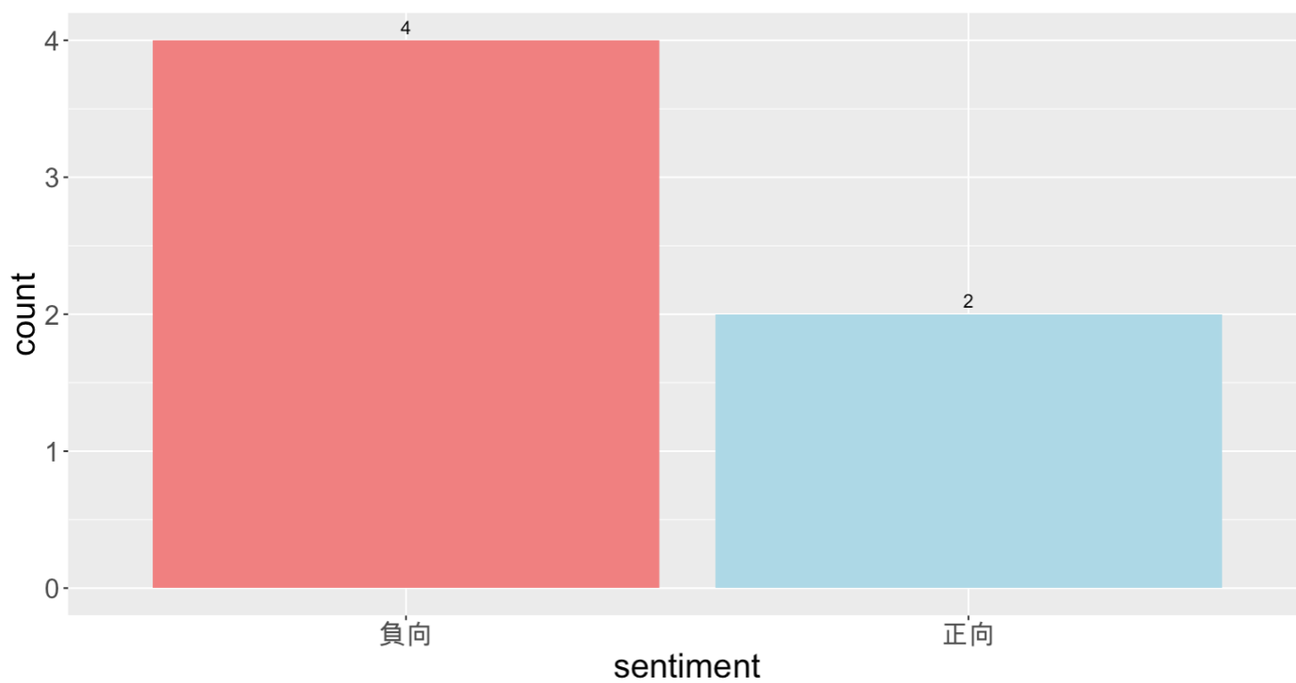
你現在的心情指標是

33

Negative WordCloud



詞頻統計



Positive WordCloud

你正在想什麼呢？

藍瘦，香菇，本來今顛高高興興，泥為什莫要說這種話？藍瘦，香菇在這裡。第一翅為一個女孩使這麼香菇，藍瘦。泥為什摸要說射種話，丟我一個人晒這裡，香菇，藍瘦在這裡，香菇...

Positive WordCloud

高興興

Negative WordCloud

瘦

告訴我

你現在的心情指標是

50

不準啦`´

我要...

請以空白相隔詞彙

輸入範例：藍瘦 香菇

增加正向詞彙

藍瘦 香菇

增加負向詞彙

你正在想什麼呢？

藍瘦，香菇，本來今顛高高興興，泥為什莫要說這種話？藍瘦，香菇在這裡。第一翅為一個女孩使這麼香菇，藍瘦。泥為什摸要說射種話，丟我一個人晒這裡，香菇，藍瘦在這裡，香菇...

Positive WordCloud

Negative WordCloud

告訴我

你現在的心情指標是

10

SIMILARITY

如何計算文本相似度？

>>> 餘弦相似性 !!!

$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\|\|B\|} = \frac{\sum_{i=1}^{n} A_i \times B_i}{\sqrt{\sum_{i=1}^{n} (A_i)^2} \times \sqrt{\sum_{i=1}^{n} (B_i)^2}}$$

A = 問題 詞頻

B = 答案 tf_idf

請選擇一個問題

如何評價洛杉磯湖人隊將答應易建聯團隊的離隊要求，裁掉易建聯？

看問題內容

請選擇一個答案

1

看答案

相似度

消息來源：Yi Jianlian asks to be released by Los Angeles Lakers重磅!曝易建聯主動要求被裁掉 因不滿在湖人角色網易體育 10 月 24 日報道：據名記 Marc Stein 報道，易建聯因為不滿意在球隊所處的地位和角色，與他的團隊已經主動向湖人要求離隊。湖人還沒有正式宣佈這一消息，最終名單將在北京時間周二公佈。隨後多家洛杉磯媒體確認了這一消息。在季前賽中，易建聯在 8 場比賽中出戰 6 次，場均打 10.7 分鐘，得到 3.0 分 2.7 籃板的數據，一共出手 20 次，命中 7 球。從季前賽可以看出，阿聯並不能在湖人確立更重要的球隊地位，而他最後也只是和托馬斯-羅賓遜以及慈世平爭奪第 15 人的位置。此前易建聯與湖人簽下的合同為 1 年 800 萬，為部分保障+獎金的合同構成：其中 25 萬為保障部分，在 2017 年 1 月 10 日如果還留在隊中，則 1139123 美元得到保障（即全額老將底薪），在此之外，當出場次數達到 20 場、40 場和 59 場時，分別可以得到 2286959 美元的觸發獎金，最高可以拿到全部的 800 萬。從目前的情況看，不但觸發激勵獎金無從談起，甚至都不一定保證能在 1 月 10 日留隊。因此易建聯很難真正算在湖人立足，主動要求離開也在情理之中。阿聯將會從湖人得到 25 萬保障部分，接下來他仍然有機會等待其他 NBA 球隊的召喚，有可能繼續著 NBA 的夢想。湖人的揭幕戰將在 10 月 27 日 10:30 主場迎戰火箭，但場上勢必不會出現易建聯的身影。<div div> 易建聯接下來該如何選擇？有哪些理想的歸宿 div>

我邏輯有問題我自己解決，但我和傻缺沒話說。分割線又來了。如果還不明白，就看看王7寫的這篇推送吧。是的，阿聯去nba是為了打球，不是為了坐冷板凳。至於來和我說湖人重建模式的，我仍然就一個觀點，湖人的路子不是最優化，別的我不想多說，因為在傳統意義的毒瘤四號位在你眼裡都是寶。也不想跟有些人廢話了，水平到沒到一打國際比賽幾個動作就很能展現出現。在濟南現場看宏遠比賽，阿聯明顯就是和cba不在一個水平線上，比之大外過之而無不及，就這個身高這個密度這個技術條件，沒有nba實力說出來誰信?時運確實有不濟，造化有時會弄人。有的時候不信，看看你自己。分割線護駕。看到最高票的回答，我不知道該說什麼。首先，我先聲明一點，我承認，阿聯離開湖人，有一部分是能力並不拔尖造成的。註意，是能力不拔尖，不是打不了。然後，我得說，阿聯離開湖人是正確的選擇，打比賽的對抗與競爭永遠比訓練更能提升一個球員。再然後，第三點，阿聯是和湖人說再見了，不是永遠告別了nba。最後一點，阿聯可能真的再也不會打nba了。可能你說了，你這不每一條都給自己留著說話的後路了嗎?那行，我說明白的，阿聯，有nba實力，但是由於性格和對抗，對環境的不適應(第一次)，由於戰術的不匹配，重新歸來的心態變化和作為之前球隊和國家隊的核心身份(第二次)這些原因，讓他與nba永遠告別了。“阿聯打不了nba”這個命題我也有考慮過，即使他曾經以第六順位的身份從杜蘭特手裡搶過當年唯一一個月新秀最佳，即使他曾經是打出過8+5，有著不錯身體底子的潛力股，但是第一次nba之旅的害怕對抗，飄在外線，戰術

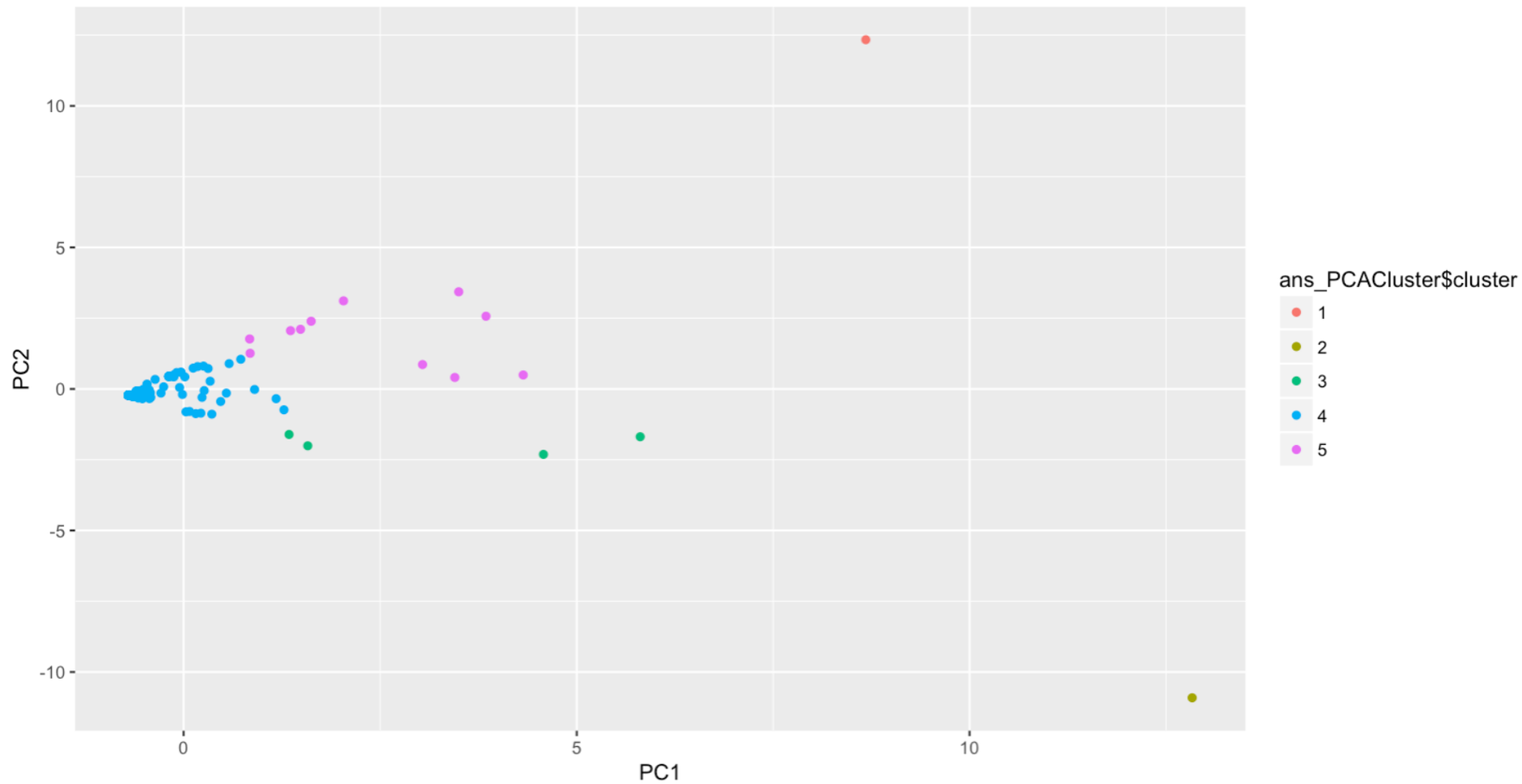
40.28 %

CLUSTERING

分群是將 Term Document Matrix

用 PCA 降階後再餵給 kmeans

下圖只有藉由 PC1 和 PC2 將分群結果部分展現



我們認為分群後
文章與各群簇的「接近度」可以視為文本的某種特徵

因此我們將此視為 features 嘗試餵給 SVM

SVM

CLUSTERING FEATURES

ESSAY SIMILARITY

SENTIMENT SCORE

SVM PROBABILITY (OAO)

STOP WORD COUNT

AUTHOR INFORMATION

WORD COUNT

CHARACTER COUNT

怎麼徹底的識別一個男人的本質？

我小姨的故事。 先直接跳到結果，她因為不能生小孩被離婚，之後受打擊太大，得了腦瘤，做了開顱手術，現在生活無法自理，思維像小孩一樣，全靠外公外婆照顧。 她年輕時很漂亮（很像舒暢），追求者眾多，但是現在連正常人都不如。而前夫娶了小三之後，生了兒子，買房買車還升了小學校長，日子過得很滋潤。 我想強調的是： 1 這個男人性格溫和有耐心，溫文爾雅。這種十多年的長期相處，性格是裝不出來的。他倆感情很好，直到離婚。我小姨比較好強，自尊心強，要面子，強勢。。 2他是入贅上門的女婿，家裡比較窮，外公家比較富裕，所以他家在那些年代受過恩惠和好處，脫了困境，曾經有很長一段時間他的親戚們都住在外公家、 3這個男人在離婚之後，我小姨開顱手術住院期間以及其後漫長的歲月里，沒探望過一次，沒任何聯繫，即使那次手術小姨差點有生命危險不管怎樣，對一個被自己傷害過的女人。怎麼這麼無情呢我無法理解。 一個看上去又顧家又有責任感的男人，卻也同時是劈腿的渣男。 以此為例，我想知道，怎樣才能看透一個人的本性，從眾多假象和偽裝中看到本質。 看評論 發現我漏了一些很重要的信息 1小姨比姨父大三歲，他倆離婚的導火索好像是異地，小姨父調任到另一個學校教書之後，不常常回來（開車得2個多小時的路程），倆人見面次數少了。他認識了一個離異的女人（我看過照片，挺漂亮)，家裡有錢，比小姨小10歲，跟前夫有孩子。後來他跟小姨就離婚了，之後又結婚生了一個兒子。 2我小時候最喜歡的就是小姨父，超過對我爸爸的喜歡，會想他是我爸爸多好。這麼多年之後，撇開小姨不談，我是懷念他的。他從來不發脾氣，帶我看影碟恐怖片，給我講腦筋急轉彎，比爸爸有耐心多了。只是現在是一個模糊的影子，長相記不大清了。 最後見他的那年春節，家裡氣氛很詭異，我當時小不知道，但突然有了隱隱的擔憂，他給我壓歲錢之後，我問他：“X叔叔，你明年還來玩嗎”媽媽瞪了我一眼，他笑著說：“會來的”。之後再沒講過話。有一次在學校遠遠的看到他 想跟他說話但沒去（我媽不讓）他看到我了 但沒跟我說話。另一次在音樂節上，遠遠看到他托舉著一個小嬰兒，笑得很開心。 3爭議最大的關於小姨的性格：她確實強勢，這是問題。但不是那麼強勢，小性子什麼的沒你們想的那麼誇張。而且她一直這樣，都知道。後來也有人追她，她沒答應。而且關健是那個小三也強勢啊管他很嚴。怪我沒說清楚。 4他的工作能力？ 我不知道 他教中學的時候我讀小學 我讀中學的時候他調走了 之後就離婚了 5家人對他的看法麼 從那以後沒有一個人提他 就像他從來不存在一樣，大家都不想揭別人傷疤。但我媽和外婆是討厭他的，外公不想談這個人，我爸爸只是說：“他胖了發福了，生了兒子，升了校長，小日子過得挺滋潤的（原話）”也不想多談。 沒人認為他比我們低一等，原來大家都很喜歡他 6有人說我來找認同感，痛批此人是渣男。沒有，大家一起罵他是渣男對事實不會有任何改變。我只是無法把我心目中的那個溫和有耐心、想讓他當我爸爸的人，和後來這個人聯繫在一起，我花了很久才相信這是同一個人。 7有人說我把小姨得腫瘤這事牽強的完全怪罪到他頭上。那麼我只陳述一個事實：離婚之後，小姨就像變了一個人，精神上消沉，具體過程我在讀書不清楚，但我媽清楚。之後的小姨跟我之前的感覺完全不一樣了，她不再像以前一樣買很多好看的衣服，不再熱衷於打扮，而且她現在的白髮比我媽還多（我媽比她大6歲）。我想強調的是，腫瘤這事不全然是離婚導致的，但必然與之脫不了關係。 8有人說我帶有濃重的主觀色彩。確實，對於一個童年我曾經最喜歡的人，我是懷念他的。對於一個拋棄我至親的小姨、傷害她最重的人，我是恨他的。你要我拋棄感情去談這個事情，我做不到。 啊哈 看完了 等到有一天找到小姨父 親自跟他聊了之後再更 人性複雜 人心變化 凡夫俗子的我們是看不透的 。所以我們都只能被命運推來搡去了麼 只能祈禱命運對我們溫柔一點了麼 你們願意認命嗎新消息 小姨要跟一個三十歲左右的退伍軍人再婚了 （年紀相差二十歲左右）感覺此人性格內向溫和 不善言辭 小姨的朋友在她結婚前夕找了前小姨父，故意告訴他小姨馬上就要跟一個各方麵條件都不錯的人結婚了，想看他反應。他第一句話是 是不是看上她的錢了？後來讓轉告他的祝福，說如果小姨過不好這輩子他都不會安心。我後來知道，這些年他自己也動過心臟手術，而且大出血，在鬼門關走了一趟。原來他內心其實有愧疚的。他過得也沒那麼好。鏡像問題：怎麼徹底的識別一個女人的本質？ - 戀愛

```r
# 用 f1 看 model 品質
the_f1_score <- F1_Score(prediction, testset$quality)
the_f1_score
```

```
[1] 0.6694915
```

請輸入你的答案

女人 豆腐心 海底針

男女以和樂為貴

還不夠如火如荼啊。
主要是我覺得目前還是誰有錢有權誰更容易風口浪尖，等有一天呂姓的家庭社會地位都高於男性了，分析這個問題會更有意義。

author_followee_num

9

author_upvote_num

27

author_thank_num

4

author_answer_num

7

author_question_num

4

author_post_num

11

看看我的回答水準

63.96

THANK YOU