## 抵制文化之現象分析

### 高嘉妤、柯堯珹、趙友誠

### **Table of contents**

資料簡介	1 2 6
抵制程度與其他因素之關聯分析 Canonical analysis and PCA-對全部變數做	6 6 8
Logistic regression model	8
Decision tree	10
glmnet	10
XGboost	11
參考文獻	13
report 的重要頁碼:	
52-網路應誘與脫序行為之子題說明	
92-資料人口結構與母群人口結構比較表	
281-各題目之測量概念	

## 資料簡介

Table 1: 變數前處理

Variables	Manipulation
q2	出生年改成年齡
q2_rr	將 rrq2 的年齡分層變數重新命名 q2_rr
q3	移除。不關心地區造成的差異
q4	重新劃分為四個等級,劃分參考人口結構表格的分類方式
q6,q7	時間統一單位(分)
q8	移除。大部分的人都有透過網路接觸名人的資訊或討論(只有四個人沒有)。
q9	移除。無法界定是工作性質或娛樂性質
q10	改成"使用幾個與 yt 名人討論相關的社群媒體",因為有些社群媒體不會造成抵制名人行為。
q11	改成"有無使用 YT 或 Twitch",原因與第十題類似。
q12~q15	移除。q28,q29 關心的時間範圍較廣並不只局限於疫情期間。
q16~q19	將每個類別補 0(變成 1,0), 再創建一個標籤變數 1719 label
q20~q26	参考碩士論文: 台灣消費者抵制行為之研究-以台商親中言論衍生之抵制為例
	(https://www.airitilibrary.com/Article/Detail/U0004-G0107932056) 之做法,將相同大主題的ordinal 主觀評分加總作為該主題程度的分數。

原始資料維度: rows×columns = 1004 × 207

Table 2: 變數解釋

Variables	Explanation	remark
ql		1: 男性, 2: 女性
q2_rr	年龄分層	1:18~29, 2:30~39, 3:40~49,
		4:50~59, 5:60~69, 6:70+
q4	教育程度	1: 高中及以下, 2: 專科,
		3: 大學, 4: 研究所
q5_1	週平均上網天數	
q6	上網分鐘 (工作、學習)	
q7	上網分鐘 (娛樂、休閒)	
q10	使用幾個與名人討論相關的社群媒體	
q11	是否使用 YT,Twitch 或 bilibili	1: 是,0: 否
q17_01	是否參與過: 不傷害、騙人	1: 是,0: 否
q17_02	是否參與過: 不傷害、不騙人	1: 是,0: 否
q19_01	是否參與過: 傷害、騙人	1: 是,0: 否
q19_02	是否參與過: 傷害、不騙人	1: 是,0: 否
q1719_label	是否至少有參與過一種網路惡搞	1: 是,0: 否
q20	主動激化 (引戰) 行為接受度	(接受)2~10(可以接受)
q22	他人攻擊行為的頻率	(從來沒有)5~20(經常)
q23	自己攻擊行為的頻率	(從來沒有)5~20(經常)
q24	媒體識讀素養	(低)5~20(高)
q25	網路論戰接受度	(低)4~20(高)
q26	不文明留言的影響力	(低)3~12(高)
q27_1	抵制意圖	(弱)1~5(強)
q28_YN	是否採取過抵制行為	1: 是, 0: 否
q28_1	採取過: 取消關注	1: 是, 0: 否
q28_2	採取過: 拒絕觀看	1: 是, 0: 否
q28_3	採取過: 在網路上留言或發文指責	1: 是, 0: 否
q29_1	抵制的原因: 歧視特定國家、種族或性別	1: 是, 0: 否
q29_2	抵制的原因: 有不同的政治意識型態或價值觀	1: 是, 0: 否
q29_3	抵制的原因: 做出不道德、不正當或不合法行	1: 是, 0: 否
• –	為	
q30_1	抵制行為的有效程度	(無效)1~5(有效)
q31_1	抵制前的同理心	(沒同理)1~4(有同理)
q32_1	抵制行為的對名人的傷害程度	(不嚴重)1~5(嚴重)
q33_1	抵制行為的對自己的重要程度	(不重要)1~5(重要)
q34_1	抵制成本	(非常少)1~5(非常多)
q35_1	抵制規模感知	(小)1~5(大)
q36_1	抵制的社會壓力	(小)1~4(大)
q38	心理幸福感	(不滿意)2~10(滿意)
q39_1	生活品質	(不快樂)1~5(快樂)
q40	國民黨喜好程度	(不喜歡)0~5 (喜歡)
q41	民進黨喜好程度	(不喜歡) 0~5 (喜歡)
q42_1	意識形態	(台獨)0~10: (統一)
weight	人口結構修正權重	(

### 敘述統計

# 40 Variables DB.csv 1004 Observations

q1	
n missing distinct Info Mean Gmd 1004 0 2 0.724 1.594 0.4829	
Value 1 2 Frequency 408 596	
Proportion 0.406 0.594 q2_rr	
n missing distinct Info Mean Gmd	
1004	
Value 7	
For the frequency table, variable is rounded to the nearest 0	
q4	, , , , , , , , , , , , , , , , , , ,
n missing distinct Info Mean Gmd 1004 0 4 0.817 2.739 0.9407	
Value 1 2 3 4 Frequency 155 121 559 169 Proportion 0.154 0.121 0.557 0.168	
For the frequency table, variable is rounded to the nearest 0	
q5	
n missing distinct Info Mean Gmd .05 .10 .25 .50 .75 .90 .95 1004 0 13 0.277 6.658 0.6393 4.0 6.5 7.0 7.0 7.0 7.0 7.0	
Value 0.5 1.0 1.5 2.0 3.0 3.5 4.0 4.5 5.0 5.5 6.0 6.5 7.0 Frequency 16 8 3 8 6 9 6 1 15 6 14 11 901	
Proportion 0.016 0.008 0.003 0.008 0.006 0.009 0.006 0.001 0.015 0.006 0.014 0.011 0.897  For the frequency table, variable is rounded to the nearest 0	
q6	երիրություն և և
n missing distinct. Info Mean Gmd 05 10 25 50 75 90 95	
1004	
q7	
•	
•	
n missing distinct Info Mean Gmd .05 .10 .25 .50 .75 .90 .95 1004 0 91 0.991 267.9 177.9 60 90 150 240 330 480 600	
n missing distinct Info Mean Gmd .05 .10 .25 .50 .75 .90 .95 1004 0 91 0.991 267.9 177.9 60 90 150 240 330 480 600 lowest: 0 7 20 30 50, highest: 900 960 1020 1035 1200 q10  n missing distinct Info Mean Gmd	
n missing distinct Info Mean Gmd .05 .10 .25 .50 .75 .90 .95 1004 0 91 0.991 267.9 177.9 60 90 150 240 330 480 600 lowest: 0 7 20 30 50, highest: 900 960 1020 1035 1200 q10  n missing distinct Info Mean Gmd 1004 0 8 0.94 2.388 1.476  Value 0 1 2 3 4 5 6 7	
n missing distinct Info Mean Gmd .05 .10 .25 .50 .75 .90 .95 1004 0 91 0.991 267.9 177.9 60 90 150 240 330 480 600 lowest: 0 7 20 30 50, highest: 900 960 1020 1035 1200 q10  n missing distinct Info Mean Gmd 1004 0 8 0.94 2.388 1.476  Value 0 1 2 3 4 5 6 7 Frequency 44 224 336 217 101 56 15 11 Proportion 0.044 0.223 0.335 0.216 0.101 0.056 0.015 0.011	
n missing distinct Info Mean Gmd .05 .10 .25 .50 .75 .90 .95 1004 0 91 0.991 267.9 177.9 60 90 150 240 330 480 600 lowest: 0 7 20 30 50, highest: 900 960 1020 1035 1200 q10  n missing distinct Info Mean Gmd 1004 0 8 0.94 2.388 1.476  Value 0 1 2 3 4 5 6 7 Frequency 44 224 336 217 101 56 15 11 Proportion 0.044 0.223 0.335 0.216 0.101 0.056 0.015 0.011  For the frequency table, variable is rounded to the nearest 0	
n missing distinct Info Mean Gmd .05 .10 .25 .50 .75 .90 .95  1004 0 91 0.991 267.9 177.9 60 90 150 240 330 480 600   lowest: 0 7 20 30 50, highest: 900 960 1020 1035 1200   q10  n missing distinct Info Mean Gmd  1004 0 8 0.94 2.388 1.476  Value 0 1 2 3 4 5 6 7  Frequency 44 224 336 217 101 56 15 11  Proportion 0.044 0.223 0.335 0.216 0.101 0.056 0.015 0.011  For the frequency table, variable is rounded to the nearest 0  q11  n missing distinct Info Mean Gmd	
n missing distinct Info Mean Gmd .05 .10 .25 .50 .75 .90 .95  1004 0 91 0.991 267.9 177.9 60 90 150 240 330 480 600   lowest: 0 7 20 30 50, highest: 900 960 1020 1035 1200   q10  n missing distinct Info Mean Gmd   1004 0 8 0.94 2.388 1.476  Value 0 1 2 3 4 5 6 7   Frequency 44 224 336 217 101 56 15 11   Proportion 0.044 0.223 0.335 0.216 0.101 0.056 0.015 0.011  For the frequency table, variable is rounded to the nearest 0  q11  n missing distinct Info Mean Gmd   1004 0 3 0.235 1.022 0.1637	
n missing distinct Info Mean Gmd .05 .10 .25 .50 .75 .90 .95  1004 0 91 0.991 267.9 177.9 60 90 150 240 330 480 600   lowest: 0 7 20 30 50, highest: 900 960 1020 1035 1200   q10  n missing distinct Info Mean Gmd  1004 0 8 0.94 2.388 1.476  Value 0 1 2 3 4 5 6 7  Frequency 44 224 336 217 101 56 15 11  Proportion 0.044 0.223 0.335 0.216 0.101 0.056 0.015 0.011  For the frequency table, variable is rounded to the nearest 0  q11  n missing distinct Info Mean Gmd	
n missing distinct Info Mean Gmd .05 .10 .25 .50 .75 .90 .95  1004 0 91 0.991 267.9 177.9 60 90 150 240 330 480 600   lowest: 0 7 20 30 50, highest: 900 960 1020 1035 1200   q10  n missing distinct Info Mean Gmd   1004 0 8 0.94 2.388 1.476   Value 0 1 2 3 4 5 6 7  Frequency 44 224 336 217 101 56 15 11  Proportion 0.044 0.223 0.335 0.216 0.101 0.056 0.015 0.011   For the frequency table, variable is rounded to the nearest 0   q11  n missing distinct Info Mean Gmd   1004 0 3 0.235 1.022 0.1637   Value 0 1 2	
n missing distinct Info Mean Gmd .05 .10 .25 .50 .75 .90 .95  lowest: 0 7 20 30 50, highest: 900 960 1020 1035 1200  q10  n missing distinct Info Mean Gmd	
n missing distinct Info Mean Gmd .05 .10 .25 .50 .75 .90 .95  1004 0 91 0.991 267.9 177.9 60 90 150 240 330 480 600   lowest: 0 7 20 30 50, highest: 900 960 1020 1035 1200   q10  n missing distinct Info Mean Gmd   1004 0 8 0.94 2.388 1.476   Value 0 1 2 3 4 5 6 7  Frequency 44 224 336 217 101 56 15 11  Proportion 0.044 0.223 0.335 0.216 0.101 0.056 0.015 0.011   For the frequency table, variable is rounded to the nearest 0   q11  n missing distinct Info Mean Gmd   1004 0 3 0.235 1.022 0.1637   Value 0 1 2	
n missing distinct Info Mean Gmd .05 .10 .25 .50 .75 .90 .95 .1004 0 91 0.991 267.9 177.9 60 90 150 240 330 480 600	
n missing distinct Info Mean Gmd .05 .10 .25 .50 .75 .90 .95  1004 0 91 0.991 267.9 177.9 60 90 150 240 330 480 600   lowest: 0 7 20 30 50, highest: 900 960 1020 1035 1200   q10  n missing distinct Info Mean Gmd	
n missing distinct Info Mean Gmd .05 .10 .25 .50 .75 .90 .95 .1004 .0 .91 .0.991 .267.9 .177.9 .60 .90 .150 .240 .330 .480 .600	
n missing distinct Info Mean Gmd .05 .10 .25 .50 .75 .90 .95   1004 0 91 0.991 267.9 177.9 60 90 150 240 330 480 600   10west: 0 7 20 30 50, highest: 900 960 1020 1035 1200    q10  n missing distinct Info Mean Gmd	
n missing distinct Info Mean Gmd .05 .10 .25 .50 .75 .90 .95 .1004 0 .91 .0.991 .267.9 177.9 60 90 150 .240 330 480 600	

q1719_label	
n missing distinct Info Sum Mean Gmd 1004 0 2 0.338 130 0.1295 0.2257	
q20	L
n missing distinct Info Mean Gmd 1004 0 9 0.785 2.925 1.33	
Value 2 3 4 5 6 7 8 9 10 Frequency 596 140 135 54 57 10 8 1 3 Proportion 0.594 0.139 0.134 0.054 0.057 0.010 0.008 0.001 0.003 For the frequency table, variable is rounded to the nearest 0	
q22	
n missing distinct Info Mean Gmd .05 .10 .25 .50 .75 .90 .95 1004 0 16 0.987 15.05 3.851 10 10 13 15 18 20 20	
Value 5 6 7 8 9 10 11 12 13 14 15 16 17 18 Frequency 10 3 6 9 12 82 54 50 70 95 193 90 52 70 Proportion 0.010 0.003 0.006 0.009 0.012 0.082 0.054 0.050 0.070 0.095 0.192 0.090 0.052 0.070	
Value 19 20 Frequency 72 136 Proportion 0.072 0.135	
For the frequency table, variable is rounded to the nearest 0	
q23	L
n missing distinct Info Mean Gmd .05 .10 .25 .50 .75 .90 .95 1004 0 14 0.92 6.989 2.504 5 5 5 6 8 10 12	
Value 5 6 7 8 9 10 11 12 13 14 15 16 17 20 Frequency 423 137 115 84 76 82 30 23 14 6 7 3 2 2 Proportion 0.421 0.136 0.115 0.084 0.076 0.082 0.030 0.023 0.014 0.006 0.007 0.003 0.002 0.002	
For the frequency table, variable is rounded to the nearest 0	
q24	
n missing distinct Info Mean Gmd .05 .10 .25 .50 .75 .90 .95 1004 0 16 0.985 13.74 2.95 9 10 12 14 15 17 18	
Value 5 6 7 8 9 10 11 12 13 14 15 16 17 18 Frequency 9 5 7 9 27 70 61 106 117 183 164 107 72 46 Proportion 0.009 0.005 0.007 0.009 0.027 0.070 0.061 0.106 0.117 0.182 0.163 0.107 0.072 0.046	
Value 19 20 Frequency 13 8 Proportion 0.013 0.008	
For the frequency table, variable is rounded to the nearest 0	
q25	Landatalara
n missing distinct Info Mean Gmd .05 .10 .25 .50 .75 .90 .95 1004 0 17 0.985 9.188 4.175 4 4 6 9 12 14 16	
Value 4 5 6 7 8 9 10 11 12 13 14 15 16 17 Frequency 162 34 80 44 160 57 102 56 163 27 43 16 30 6 Proportion 0.161 0.034 0.080 0.044 0.159 0.057 0.102 0.056 0.162 0.027 0.043 0.016 0.030 0.006	
Value 18 19 20 Frequency 9 3 12 Proportion 0.009 0.003 0.012	
For the frequency table, variable is rounded to the nearest 0	
q26	
n missing distinct Info Mean Gmd .05 .10 .25 .50 .75 .90 .95 1004 0 10 0.941 9.47 2.536 5 6 9 9 12 12 12	
Value 3 4 5 6 7 8 9 10 11 12 Frequency 42 6 6 84 38 47 307 100 69 305 Proportion 0.042 0.006 0.006 0.084 0.038 0.047 0.306 0.100 0.069 0.304	
For the frequency table, variable is rounded to the nearest 0	
q27	. 1 1 1 .
n missing distinct Info Mean Gmd 1004 0 5 0.925 3.102 1.181	
Value 1 2 3 4 5 Frequency 73 209 368 251 103 Proportion 0.073 0.208 0.367 0.250 0.103	
For the frequency table, variable is rounded to the nearest 0	
q28_YN	
n missing distinct Info Sum Mean Gmd 1004 0 2 0.63 703 0.7002 0.4203	

q28_1	
n missing distinct Info Sum Mean Gmd 1004 0 2 0.75 490 0.488 0.5002	
q28_2	
n missing distinct Info Sum Mean Gmd 1004 0 2 0.726 591 0.5886 0.4848	
q28_3	
n missing distinct Info Sum Mean Gmd 1004 0 2 0.155 55 0.05478 0.1037	
q29_1	
n missing distinct Info Sum Mean Gmd 1004 0 2 0.706 381 0.3795 0.4714	
q29_2	
n missing distinct Info Sum Mean Gmd 1004 0 2 0.62 293 0.2918 0.4137	
q29_3	
n missing distinct Info Sum Mean Gmd 1004 0 2 0.723 598 0.5956 0.4822	
q30	l l l .
n missing distinct Info Mean Gmd 1004 0 6 0.936 2.299 1.896	
Value 0 1 2 3 4 5 Frequency 301 45 90 235 287 46 Proportion 0.300 0.045 0.090 0.234 0.286 0.046	
For the frequency table, variable is rounded to the nearest 0	
q31	1 , 1 1 ,
n missing distinct Info Mean Gmd 1004 0 5 0.924 1.784 1.491	
Value 0 1 2 3 4 Frequency 301 80 222 337 64 Proportion 0.300 0.080 0.221 0.336 0.064	
For the frequency table, variable is rounded to the nearest 0	
q32	1
n missing distinct Info Mean Gmd 1004 0 6 0.927 2.453 1.921	
Value 0 1 2 3 4 5 Frequency 301 14 72 227 326 64 Proportion 0.300 0.014 0.072 0.226 0.325 0.064	
For the frequency table, variable is rounded to the nearest 0	
q33	Landa Landa
n missing distinct Info Mean Gmd 1004 0 6 0.932 2.017 1.695	
Value 0 1 2 3 4 5 Frequency 301 57 155 328 141 22 Proportion 0.300 0.057 0.154 0.327 0.140 0.022	
For the frequency table, variable is rounded to the nearest 0	
q34	
n missing distinct Info Mean Gmd 1004 0 6 0.925 1.429 1.372	
Value 0 1 2 3 4 5 Frequency 301 297 105 279 19 3 Proportion 0.300 0.296 0.105 0.278 0.019 0.003	
For the frequency table, variable is rounded to the nearest 0	
q35	I i i I i .
n missing distinct Info Mean Gmd 1004 0 6 0.932 1.993 1.782	
Value 0 1 2 3 4 5 Frequency 301 132 63 330 137 41 Proportion 0.300 0.131 0.063 0.329 0.136 0.041	
For the frequency table, variable is rounded to the nearest 0	

```
q36
                             Info
0.924
                   distinct
 1004
Value
Frequency
              301
                     \begin{array}{c} 1 \\ 244 \end{array}
                            320
                                 3
129
Proportion 0.300 0.243 0.319 0.128 0.010
For the frequency table, variable is rounded to the nearest {\tt 0}
                                                                                                                       distinct
                             Info
0.951
 n
1004
         missing
Value
For the frequency table, variable is rounded to the nearest {\bf 0}
q39
         missing
                             Info
0.863
                   distinct
                                     Mean
3.26
 1004
Value 1 2 3 4 5 Frequency 28 127 443 368 38 Proportion 0.028 0.126 0.441 0.367 0.038
For the frequency table, variable is rounded to the nearest 0
         missing
                   distinct
                              Info
 1004
                             0.916
                    2
171
                            3
335
                                    99
              346
Proportion 0.345 0.170 0.334 0.099 0.053
For the frequency table, variable is rounded to the nearest {\tt 0}
         missing
                             Info
0.923
 n
1004
                   distinct
Value
Frequency
                     2
167
                            3
351
                                   131
Proportion 0.300 0.166 0.350 0.130 0.054
For the frequency table, variable is rounded to the nearest {\tt 0}
q42
                             Info
0.859
                                     Mean
3.869
 n
1004
Value
                                   3
76
                                           4
82
                             53
Proportion 0.139 0.060 0.053 0.076 0.082 0.516 0.028 0.015 0.009 0.003 0.020
For the frequency table, variable is rounded to the nearest 0
weight
                                                      .10
0.2592
                                               .05
0.2251
 1004
                                                              .25 .50
0.3170 0.4422
                                                                                 0.7359
lowest: 0.159109 0.163467 0.163625 0.170256 0.176108, highest: 12.5196 13.2911 13.8745 14.1467 20.0266
```

### 各變數依有無抵制行為分類畫比例圖

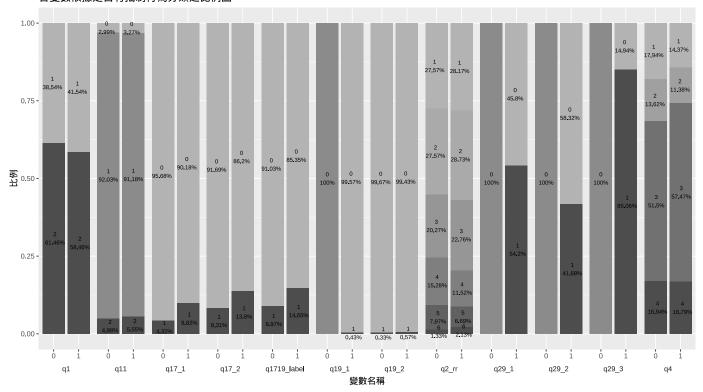
### 抵制程度與其他因素之關聯分析

#### Canonical analysis and PCA-對全部變數做

#### [1] 0.5561838 0.3779913 0.2467229

	[,1]	[,2]
q2_rr	-0.34266949	-0.12711767
q4	0.38487966	-0.15842647
q5	0.19914117	0.23277952
q6	0.26345928	0.11129500
q7	0.25597871	0.19167914
q10	0.34336636	-0.05520667
a11	0.26924437	-0.10191258

#### 各變數依據是否有抵制行為分類之比例圖



q1719_label	0.26642705	-0.11947314
q20	0.12905657	-0.31645562
q22	0.58027872	-0.04811900
q23	0.30722923	-0.21727920
q24	0.45986052	-0.02083631
q25	0.13248128	-0.11773570
q26	0.54426288	0.03699984
q27	0.44448715	-0.44852953
q28_1_2	0.02414203	0.24982439
q28_3	0.07609393	-0.35184289
q29_1	0.33046972	-0.12221603
q29_2	-0.17379321	-0.18147466
q29_3	0.31538893	0.02180216
q31	0.32667211	0.44132060
q33	0.54874451	-0.07820923
q34	0.11960599	0.33788881
q36	0.36602606	-0.02013521
q29_1_2_inter	0.13162928	-0.22028703
q29_1_3_inter	0.39016731	-0.09069338
q38	0.06768219	0.11086560
q39	0.07310548	0.15654272
q40	-0.34264406	0.26799181
q41	0.08027230	0.09291283
q42	-0.35131883	0.14988977
[,1]	[,2]	
q30 0.4995686		

[1] 0.7751 0.3213

q32 0.3064858 0.88958246 q35 0.9630194 -0.11880929

[1] 0.3177 0.2017

- [1] 0.2397 0.0459
- [1] 0.0983 0.0288

把相關性 < 0.2 的删除

#### Canonical analysis and PCA-對部分變數做

[1] 0.5493481 0.3077433 0.2154761

```
[,1]
                                 [,2]
              -0.3408722 0.116294002
q2_rr
q4
               0.3939827 0.168113039
q6
               0.2627364 -0.144760731
q7
               0.2526976 -0.241792023
               0.3475554 0.091713828
q10
q11
               0.2747058 0.129625641
               0.2725526 0.145531706
q1719_label
               0.5880407 0.025256918
q22
q23
               0.3165768 0.270230718
               0.4658894 -0.015795455
q24
q26
               0.5488353 -0.069505664
q27
               0.4639449 0.488644417
q29_1
               0.3375125 0.140158266
q29_2
             -0.1707743 0.254975893
q29_3
               0.3194057 -0.090509080
               0.3182456 -0.612910391
q31
q33
               0.5578086 0.037667626
q36
               0.3708340 -0.004645566
q29_1_2_inter 0.1395356 0.270659736
q29_1_3_inter 0.3971705 0.088206845
              -0.3548512 -0.296692495
q40
              -0.3592300 -0.180111853
q42
         [,1]
                    [,2]
q30 0.5075172 -0.1276347
q32 0.2689797 -0.9450105
q35 0.9654139 0.1148991
[1] 0.7704 0.2891
[1] 0.3155 0.2306
[1] 0.2325 0.0274
[1] 0.0952 0.0218
```

## Logistic regression model

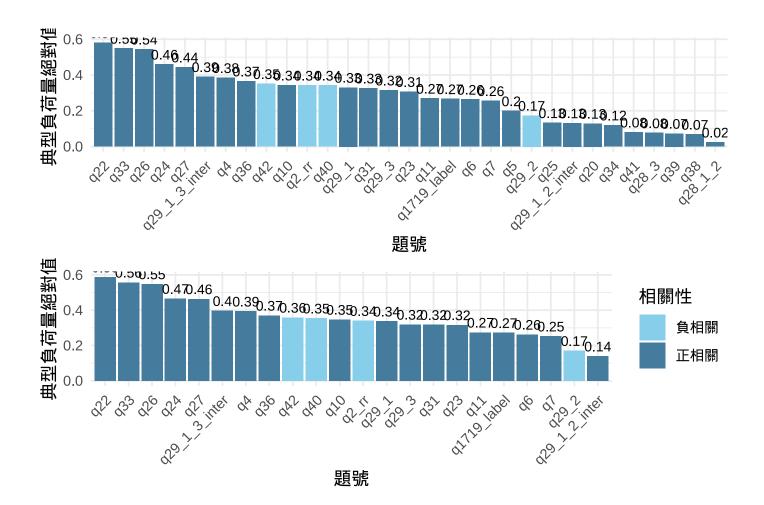
可以知道有使用 youtube 和 twitch 的人、越不能接受別人因為一些因素而罵他的人做出抵制行為的機率越小,越常做出網路攻擊行為和看到別人的攻擊行為、越想抵制名人的話就越有可能做出抵制行為。

q10 跟 q24 注意一下可能有關連

Logistic model 的變數選取方法參照Variable selection – A review and recommendations for the practicing statistician整理出的建議標準。

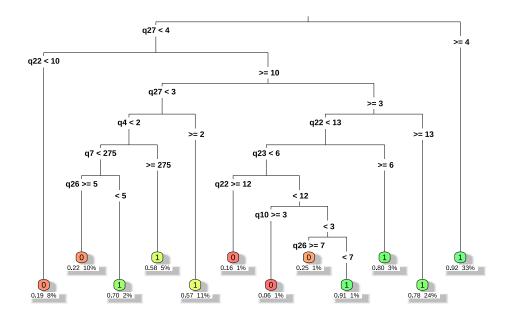
我們的 EPV(Event Per Variable) 是以 q28\_YN 作為反應變數的情況下,可使用的解釋變數是 36=41-weight-q28\_1~q28\_3-q28\_YN(反應變數)=

```
Estimate Std. Error z value Pr(>|z|)
(Intercept) -4.133369543 0.6755910429 -6.118153 9.466600e-10
q2_rr 0.138545731 0.0619525965 2.236318 2.533092e-02
q4 0.112791142 0.0797498661 1.414311 1.572705e-01
```



```
0.001077181 0.0005208604 2.068080 3.863250e-02
q7
q11
            -0.408117557 0.2667164164 -1.530155 1.259783e-01
             0.132368312 0.0247272204 5.353142 8.644015e-08
q22
q23
             0.107733712 0.0410475737 2.624606 8.674926e-03
            -0.066376973 0.0303476804 -2.187217 2.872666e-02
q24
q27
             0.942418746 0.0881405693 10.692224 1.106840e-26
                 Estimate
                            Std. Error
                                          z value
(Intercept) -4.226861e+00 0.7696724491 -5.4917666 3.979331e-08
factor(q1)2 -4.920600e-02 0.1622429723 -0.3032859 7.616720e-01
             1.220499e-01 0.0671322562
                                       1.8180518 6.905621e-02
q2_rr
q4
             1.161110e-01 0.0841034557
                                        1.3805739 1.674100e-01
                                        0.8427229 3.993835e-01
             3.819072e-02 0.0453182474
q5
q6
            -7.395442e-05 0.0004360273 -0.1696096 8.653172e-01
             1.083682e-03 0.0005427127
                                       1.9967877 4.584825e-02
q7
            -6.655493e-02 0.0687609012 -0.9679182 3.330852e-01
q10
            -4.030700e-01 0.2694425081 -1.4959406 1.346691e-01
q11
q20
             2.207562e-02 0.0626781644
                                        0.3522059 7.246838e-01
             1.342025e-01 0.0297319289 4.5137518 6.369076e-06
q22
             1.034683e-01 0.0426483262 2.4260819 1.526283e-02
q23
            -6.529589e-02 0.0313880877 -2.0802761 3.750022e-02
q24
             6.690675e-03 0.0218581909 0.3060947 7.595326e-01
q25
q26
            -7.137902e-03 0.0369011421 -0.1934331 8.466198e-01
q27
             9.411282e-01 0.0886444113 10.6168928 2.487022e-26
```

這個模型的節具象



### **Decision tree**

### glmnet

```
16 x 1 sparse Matrix of class "dgCMatrix"
(Intercept) -0.004137663
factor(q1)2
q2_rr
             0.003954779
q4
q5
             0.007744427
q6
q7
q10
q11
            -0.035702073
             0.003489347
q20
q22
             0.009061882
q23
             0.017838738
             0.003400583
q24
q25
            -0.006970239
q26
             0.137817452
q27
Call:
glm(formula = factor(q28_YN) \sim q2_rr + q4 + q5 + q6 + q11 + q20 +
    q22 + q23 + q24 + q25 + q26 + q27, family = binomial, data = DB.csv,
    weights = weight)
Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -4.181e+00 7.149e-01 -5.848 4.96e-09 ***
             1.218e-01 6.196e-02
                                    1.966
                                             0.0493 *
q2_rr
q4
             9.393e-02 8.278e-02
                                     1.135
                                             0.2565
             4.557e-02 4.449e-02
                                             0.3057
q5
                                     1.024
q6
             8.438e-05 4.169e-04
                                    0.202
                                             0.8396
q11
            -4.203e-01 2.671e-01 -1.574
                                             0.1155
```

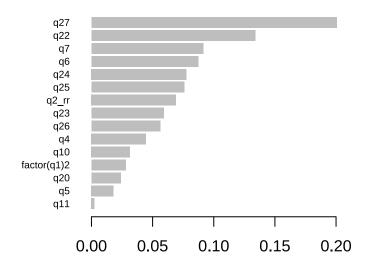
```
3.092e-02 6.207e-02 0.498 0.6184
q20
q22
           1.347e-01 2.959e-02 4.553 5.30e-06 ***
           1.067e-01 4.168e-02 2.559 0.0105 *
q23
          -7.089e-02 3.108e-02 -2.281 0.0226 *
q24
q25
           6.793e-03 2.157e-02 0.315 0.7529
q26
            2.383e-04 3.668e-02 0.006 0.9948
            9.384e-01 8.841e-02 10.615 < 2e-16 ***
q27
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Dispersion parameter for binomial family taken to be 1)
   Null deviance: 1261.7 on 1003 degrees of freedom
Residual deviance: 1028.4 on 991 degrees of freedom
AIC: 932.88
```

Number of Fisher Scoring iterations: 5

### **XGboost**

Joost				
Feature	Gain	Cover	Frequency	7
<char></char>	<num></num>	<num></num>	<num></num>	>
q27	0.200510764	0.14986138	0.056331007	7
q22	0.133835425	0.10788980	0.105184447	7
<b>q</b> 7	0.091330404	0.11710668	0.126121635	5
q6	0.087438931	0.11404843	0.122133599	)
q24	0.077734713	0.08543438	0.099700897	7
q25	0.076031973	0.07654591	0.094217348	3
q2_rr	0.069100679	0.04840269	0.063808574	ŀ
q23	0.059092613	0.08103531	0.079760718	3
q26	0.056171096	0.06543503	0.072283151	L
q4	0.044628098	0.04126913	0.055832502	2
q10	0.031542328	0.03911213	0.050847458	3
factor(q1)2	0.027981144	0.01446209	0.024925224	ŀ
q20	0.024096527	0.03192305	0.029411765	5
q5	0.018035415	0.01539175	0.013958126	3
q11	0.002469889	0.01208223	0.005483549	)
ap_score				
_	)2 q2_1	rr	q4	
<nur< td=""><td>n&gt; <nur< td=""><td>r&gt; <n< td=""><td>num&gt;</td><td><nu< td=""></nu<></td></n<></td></nur<></td></nur<>	n> <nur< td=""><td>r&gt; <n< td=""><td>num&gt;</td><td><nu< td=""></nu<></td></n<></td></nur<>	r> <n< td=""><td>num&gt;</td><td><nu< td=""></nu<></td></n<>	num>	<nu< td=""></nu<>
	Feature	Feature Gain	Feature Gain Cover <char> <num> <num> <num> q27 0.200510764 0.14986138 q22 0.133835425 0.10788980 q7 0.091330404 0.11710668 q6 0.087438931 0.11404843 q24 0.077734713 0.08543438 q25 0.076031973 0.07654591 q2_rr 0.069100679 0.04840269 q23 0.059092613 0.08103531 q26 0.056171096 0.06543503 q4 0.044628098 0.04126913 q10 0.031542328 0.03911213 factor(q1)2 0.027981144 0.01446209 q20 0.024096527 0.03192305 q5 0.018035415 0.01539175 q11 0.002469889 0.01208223 ap_score factor(q1)2 q2_rr</num></num></num></char>	Feature Gain Cover Frequency <char> <num> <num <="" <num="" <num<="" num="" td=""></num></num></num></num></num></num></num></num></num></num></num></num></num></num></num></num></num></num></num></num></num></num></num></num></num></num></num></num></num></num></num></num></num></num></num></num></num></num></num></num></num></num></num></num></num></char>

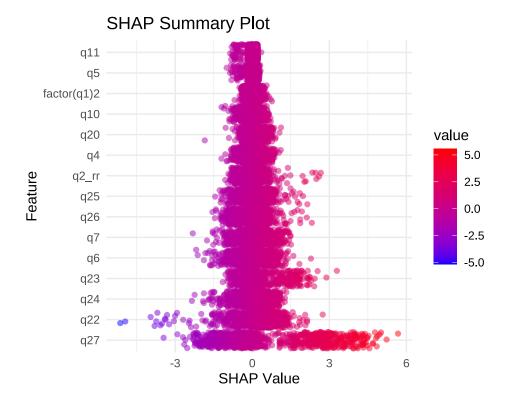
					_score	фыцар,
q7	q6	q5	q4	q2_rr	factor(q1)2	
<num></num>	<num></num>	<num></num>	<num></num>	<num></num>	<num></num>	
0.61428833	-0.03501144	0.107510969	-0.514282167	0.15691492	-0.05291108	1:
-0.07906581	-0.25388312	0.044495411	0.500798821	0.66198993	-0.03804322	2:
-0.66021556	0.91300434	0.049778830	-0.217827514	0.10885221	0.30124280	3:
0.21622194	-0.35678056	-0.013059865	0.007799798	-0.07434784	-0.15189847	4:
-0.07468783	0.07656764	0.222510144	-0.349931628	0.20496246	-0.05766268	5:
-0.11328854	-0.58729368	-0.003203744	0.013911518	-0.61297596	0.04110148	1000:
0.92444080	0.72794259	0.026580764	-0.651927292	-0.15809917	0.15003079	1001:
0.49518225	0.31439698	0.002380930	0.026317565	-0.32781160	-0.08099592	1002:
0.04893903	-0.35419035	0.049558733	0.298285365	-0.08020609	-0.02017495	1003:
0.21641597	0.46798557	0.037077535	-0.175045028	0.23888724	0.06181900	1004:
q24	q23	q22	q20	q11	q10	
<num></num>	<num></num>	<num></num>	<num></num>	> <num></num>	<num< td=""><td></td></num<>	
0.14522108	-0.243571937	-1.34608972	-0.254262596	0.01783300	0.12745729	1:
-0.05042689	1.800945044	0.52687192	0.003492521	9 0.02474033	0.036662769	2:
-0.11986512	-0.120019943	1.25518298	0.044680826	0.01991163	-0.187671736	3:
-0.47970963	-0.091351420	0.60540199	0.127911299	0.01820237	0.20190519	4:
-0.24068443	1.542876005	0.31676680	-0.143812791	0.01309785	0.083755620	5:



```
1000: 0.154443979 0.01938938 -0.174130142 -0.19928446 -0.061704461
                                                                  0.69123656
1001: 0.006108247 0.01923613 -0.084150225
                                          0.72979736 -0.051257744
                                                                  0.57152742
1002: -0.206083208 0.01426584 -0.143989041
                                          0.41245130 0.461106896
                                                                  0.72220165
1003: 0.189425349 0.02641905 0.439815134
                                          0.06346079 -0.328560501
                                                                  0.64774746
1004: -0.252529025 0.01957053 -0.015451639
                                          0.69806492 0.008263929
                                                                  0.70546347
             q25
                         q26
                                    q27
           <num>
                       <num>
  1: -0.05801377 -0.16845839 -0.57342768
      0.48889813 0.28977826 0.37257659
  4: -0.70940357 -0.60682899 -1.00455034
      0.18101855 -0.40762818 3.62704206
  5:
1000: 0.06543855 -0.59360772 -1.61921859
1001:
      0.16276813 -0.61529320 -0.09688544
1002:
      1.36524403 -0.50274253 -0.67414463
1003: 1.84676373 0.05711459 -1.33648241
1004: 0.65931511 0.19910364 3.51921916
$mean_shap_score
                   q22
                               q24
                                          q23
       q27
                                                       q6
1.45027136
            0.61723131
                       0.46055746
                                   0.44620832
                                               0.42553986
                                                                  q10
       q26
                   q25
                             q2_rr
                                           q4
                                                      q20
0.40501578
            0.32988680
                        0.27816917
                                   0.25791766
                                              0.19946306
                                                          0.18716857
                              q11
factor(q1)2
                    q5
0.13234583 0.08041424 0.04918274
$BIASO
```

BIAS <num>

1: 0.9782695



## 參考文獻

- [1] 台灣消費者抵制行為之研究—以台商親中言論衍生之抵制為例
- [2] Variable selection A review and recommendations for the practicing statistician

