

抵制文化之現象分析

高嘉好、柯堯城、趙友誠

Table of contents

Memo	1
資料簡介與資料視覺化	1
敘述統計	2
各變數依有無抵制行為分類畫比例圖	6
分析影響抵制與否之因素	7
Logistic regression model	7
glmnet	8
Decision tree	8
XGboost	8
分析影響抵制程度之因素 (典型相關分析 CCA)	8
Canonical Correlation Analysis and PCA-對全部變數做	9
Canonical analysis and PCA-對部分變數做	9
參考文獻	11

Memo

report 的重要頁碼: 52-網路癮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

資料簡介與資料視覺化

對資料各變數之前處理如表 1 所示:

Table 1: 變數前處理

變數	處理
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	參考碩士論文: 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

處理後資料維度: rows×columns = 1004 × 40

Table 2: 變數解釋

變數名稱	解釋	備註
q1	性別	1: 男性, 2: 女性
q2_rr	年齡分層	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	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	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	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、不正當或不合法行為	1: 是, 0: 否
q30	抵制行為的有效程度	(無效)1~5(有效)
q31	抵制前的同理心	(沒同理)1~4(有同理)
q32	抵制行為的對名人的傷害程度	(不嚴重)1~5(嚴重)
q33	抵制行為的對自己的重要程度	(不重要)1~5(重要)
q34	抵制成本	(非常少)1~5(非常多)
q35	抵制規模感知	(小)1~5(大)
q36	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	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	意識形態	(台獨)0~10: (統一)
weight	人口結構修正權重	

敘述統計

以 Hmisc::describe() 的方式繪製各變數之直方圖與敘述統計量。

DB.csv					
40 Variables				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	0	6	0.942	2.481	1.436

Value	1	2	3	4	5	6
Frequency	281	285	221	127	71	19
Proportion	0.280	0.284	0.220	0.126	0.071	0.019

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

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1004	0	92	0.995	264.6	239.6	0	0	90	240	420	540	600

lowest : 0 1 5 10 15, highest: 900 960 1080 1200 1440

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

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
Frequency	32	918	54
Proportion	0.032	0.914	0.054

For the frequency table, variable is rounded to the nearest 0

q17_1

n	missing	distinct	Info	Sum	Mean	Gmd
1004	0	2	0.225	82	0.08167	0.1502

q17_2

n	missing	distinct	Info	Sum	Mean	Gmd
1004	0	2	0.32	122	0.1215	0.2137

q19_1

n	missing	distinct	Info	Sum	Mean	Gmd
1004	0	2	0.009	3	0.002988	0.005964

q19_2

n	missing	distinct	Info	Sum	Mean	Gmd
1004	0	2	0.015	5	0.00498	0.00992

q1719

n	missing	distinct	Info	Sum	Mean	Gmd
1004	0	2	0.338	130	0.1295	0.2257

q20

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

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

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

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

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

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

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

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

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

n	missing	distinct	Info	Mean	Gmd
1004	0	5	0.924	1.306	1.176

Value	0	1	2	3	4
Frequency	301	244	320	129	10
Proportion	0.300	0.243	0.319	0.128	0.010

For the frequency table, variable is rounded to the nearest 0

q38

n	missing	distinct	Info	Mean	Gmd
1004	0	9	0.951	6.232	1.561

Value	2	3	4	5	6	7	8	9	10
Frequency	12	22	81	154	301	247	151	30	6
Proportion	0.012	0.022	0.081	0.153	0.300	0.246	0.150	0.030	0.006

For the frequency table, variable is rounded to the nearest 0

q39

n	missing	distinct	Info	Mean	Gmd
1004	0	5	0.863	3.26	0.8707

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

q40

n	missing	distinct	Info	Mean	Gmd
1004	0	5	0.916	2.345	1.31

Value	1	2	3	4	5
Frequency	346	171	335	99	53
Proportion	0.345	0.170	0.334	0.099	0.053

For the frequency table, variable is rounded to the nearest 0

q41

n	missing	distinct	Info	Mean	Gmd
1004	0	5	0.923	2.472	1.321

Value	1	2	3	4	5
Frequency	301	167	351	131	54
Proportion	0.300	0.166	0.350	0.130	0.054

For the frequency table, variable is rounded to the nearest 0

q42

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1004	0	11	0.859	3.869	2.235	0	0	2	5	5	5	6

Value	0	1	2	3	4	5	6	7	8	9	10
Frequency	140	60	53	76	82	518	28	15	9	3	20
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

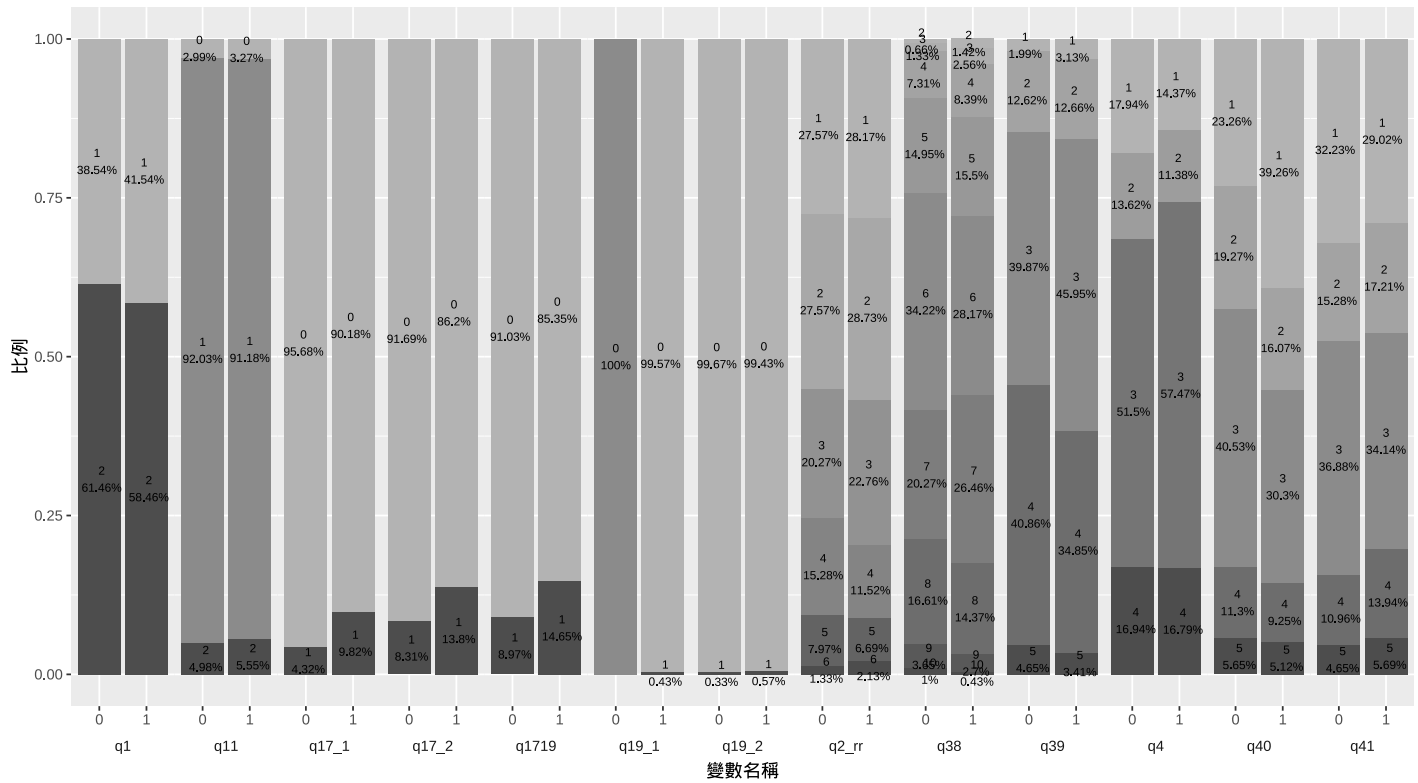
n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1004	0	387	1	1.001	1.122	0.2251	0.2592	0.3170	0.4422	0.7359	2.7754	3.8788

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

以抵制與否作為依據，將各變數的類別比例繪製出來。其中並無繪製類別過多或是 q28 以後的變數。

各變數依據是否有抵制行為分類之比例圖



可以發現有些變數的比例十分不平衡，如 q19，預期在進行變數挑選或是分析時不會起到太大作用。除此之外並無發現明顯比例差異。

分析影響抵制與否之因素

Logistic regression model

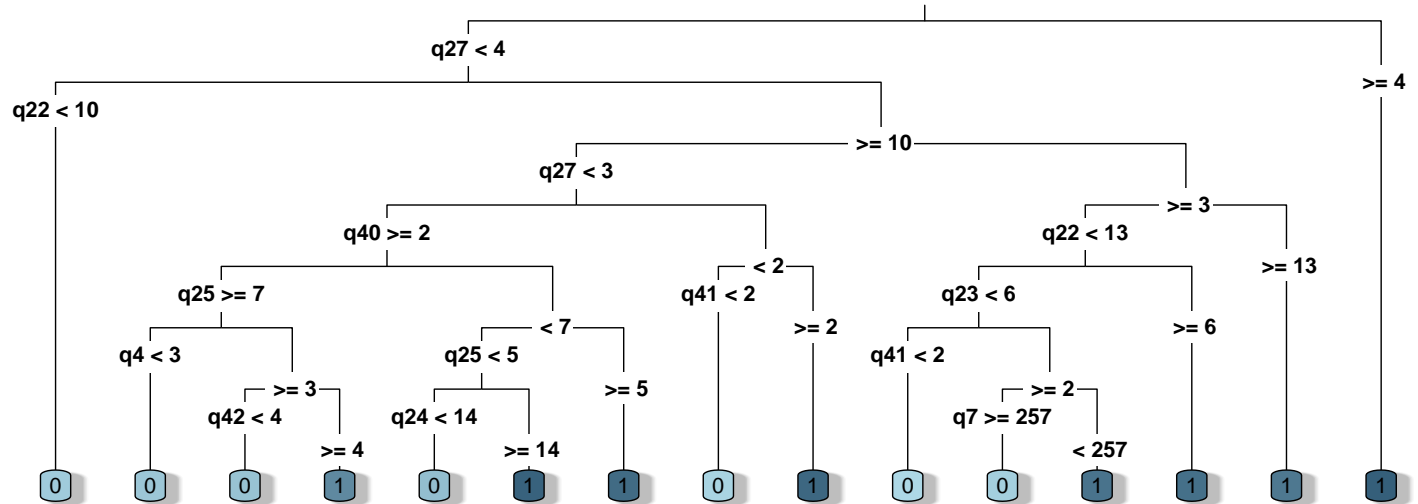
我們採用逐步選取法 (Heinze, 2018) 對可能影響抵制行為的變數進行篩選並加上人口權重。結果如表 2 所示，此模型有 10 個顯著變數，以下針對顯著變數解讀：1. 抵制意圖 (q27 數值評分:1~5)：抵制意圖越高的人，更容易付諸實際行動。每增加 1 單位使得抵制行為勝算比變成 2.56 倍。2. 民進黨好感度 (q41 數值評分:1~5)：對民進黨有好感，傾向採取抵制行動。每增加 1 單位使得勝算比變為 1.36 倍。3. 年齡分層 (q2_rr 數值分層:1~6)：年齡越大的人，越可能有抵制行動。每往上移動一層，勝算比變成 1.23 倍。4. 教育程度 (q4 數值評分:1~4)：教育程度越高，採取抵制行動的可能性越大。例如研究所學歷對大學學歷的勝算比是 1.18。5. 看見網路攻擊頻率與自己網路攻擊頻率 (q22、q23 數值評分:5~20)：越常接觸他人的負面言論，或自身越常發表負面言論，均會顯著提高抵制的可能性。每增加一單位分別使得勝算比變為 1.13 及 1.11 倍。6. 休閒娛樂上網時間 (q7 分鐘)：上網時間越長，抵制行為的可能性越高，呈現正相關，每多半小時勝算比變成 1.03 倍。3 7. 意識形態 (q42 評分:1~10)：獨立意識形態 (評分較低) 的人，傾向採取抵制行動。評分每增加 1 單位使得勝算比變成 0.92 倍。8. 心理幸福感 (q38 評分:2~10)：較快樂 (評分較高) 的人，較少採取抵制行動。評分每增加 1 單位，勝算比變成 0.75 倍。9. 使用可討論名人之平台數量 (q11 計數:0,1,2)：用越多的人，越少採取抵制行動。每多用一個，勝算比變成 0.59 倍。

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-2.899384399	0.7993167194	-3.627329	2.863687e-04
q2_rr	0.203587963	0.0652825135	3.118568	1.817322e-03
q4	0.164989040	0.0820190357	2.011594	4.426270e-02
q7	0.001100082	0.0005324809	2.065957	3.883257e-02
q11	-0.534606108	0.2719071417	-1.966135	4.928304e-02
q22	0.126316011	0.0252948510	4.993744	5.921987e-07
q23	0.107438911	0.0418442591	2.567590	1.024081e-02
q24	-0.059494898	0.0308356307	-1.929421	5.367868e-02
q27	0.940311361	0.0903335051	10.409331	2.247956e-25
q38	-0.284900115	0.0666189418	-4.276563	1.898007e-05
q41	0.303523628	0.0744094856	4.079099	4.521066e-05
q42	-0.086846731	0.0394086351	-2.203749	2.754202e-02

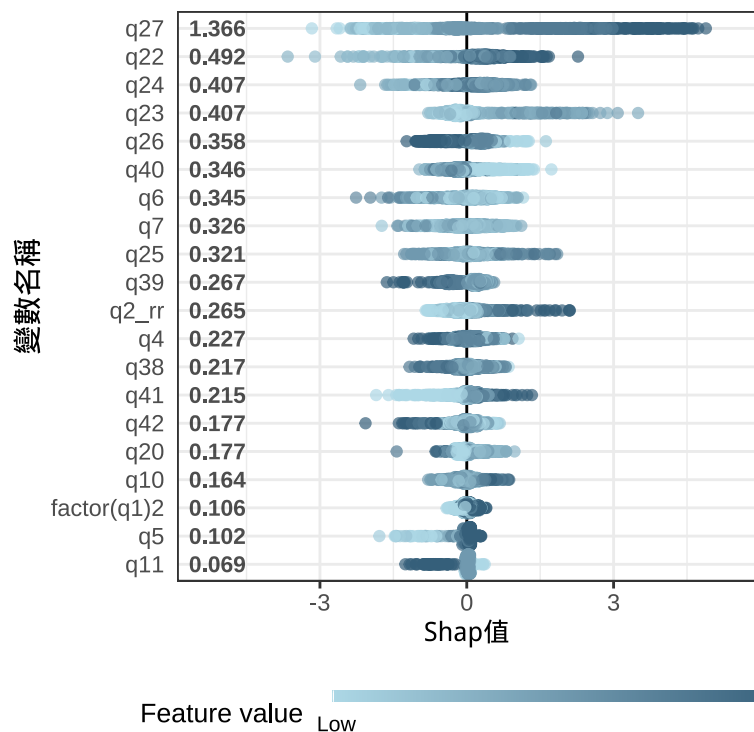
glmnet

Decision tree

決策樹選出的重要變數和邏輯式模型大致相同，但更能捕捉變數間交互作用。在圖 3 中，先從抵制意願做區分，具有高度抵制意願的人，會有抵制行為。然而較不具備抵制意願的人會受他人網路攻擊行為影響，如果常看到他人網路攻擊行為，便可能有抵制行為；反之則不會有抵制行為。另外，就算經常看到他人攻擊言論，如果對國民黨好感度較高、能接受網路攻擊行為且教育程度較低，也不會有抵制行為。



XGboost



分析影響抵制程度之因素 (典型相關分析 CCA)

我們認為抵制程度能用「抵制效度」、「抵制傷害程度」、「抵制規模」這三個問題來衡量，因此我們將這些變數設為反應變數、其他的變數設為解釋變數，去進行典型相關分析。

Canonical Correlation Analysis and PCA-對全部變數做

典型相關係數 0.5561838 0.3779913 0.2467229

解釋變數之自我相關係數 0.7751 0.3213

反應變數之自我相關係數 0.3177 0.2017

[1] 0.2397 0.0459

[1] 0.0983 0.0288

分析結果發現: 典型相關係數為 0.56、0.38、0.25。但是因為有 11 個變數的值和第一典型變數的相關性小於 0.2，因此決定將這些變數去除後，再做一次典型相關分析。

Canonical analysis and PCA-對部分變數做

這次的結果顯示: 典型相關係數為 0.55、0.31、0.22。

第一典型變數主要由

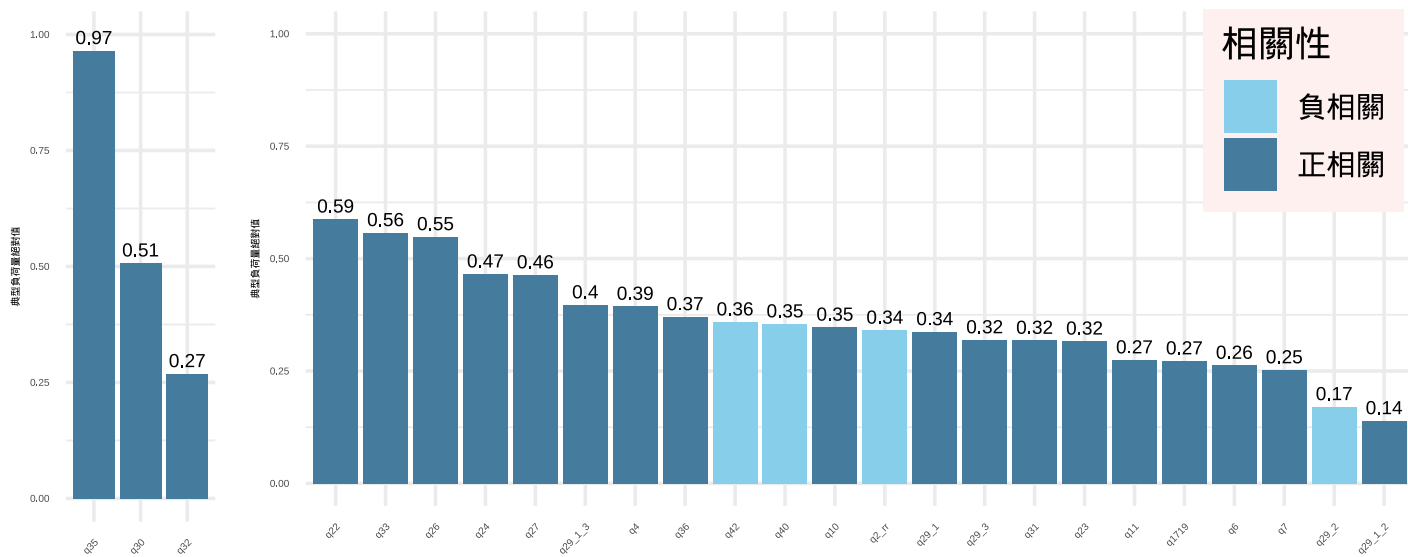
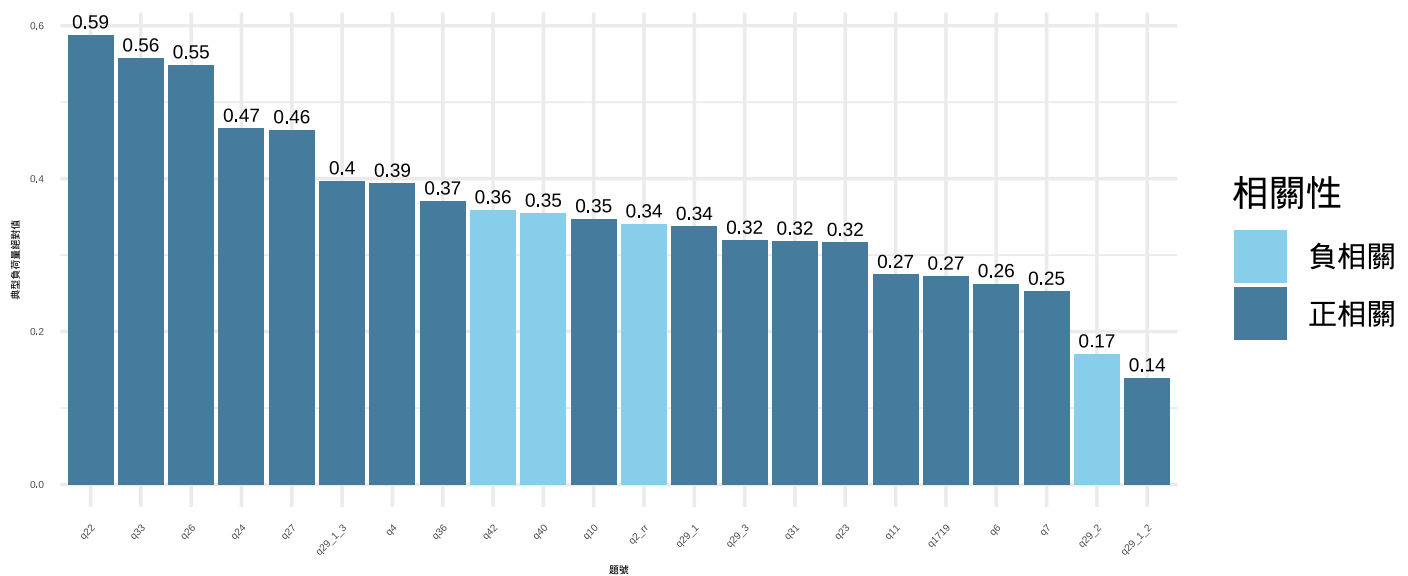
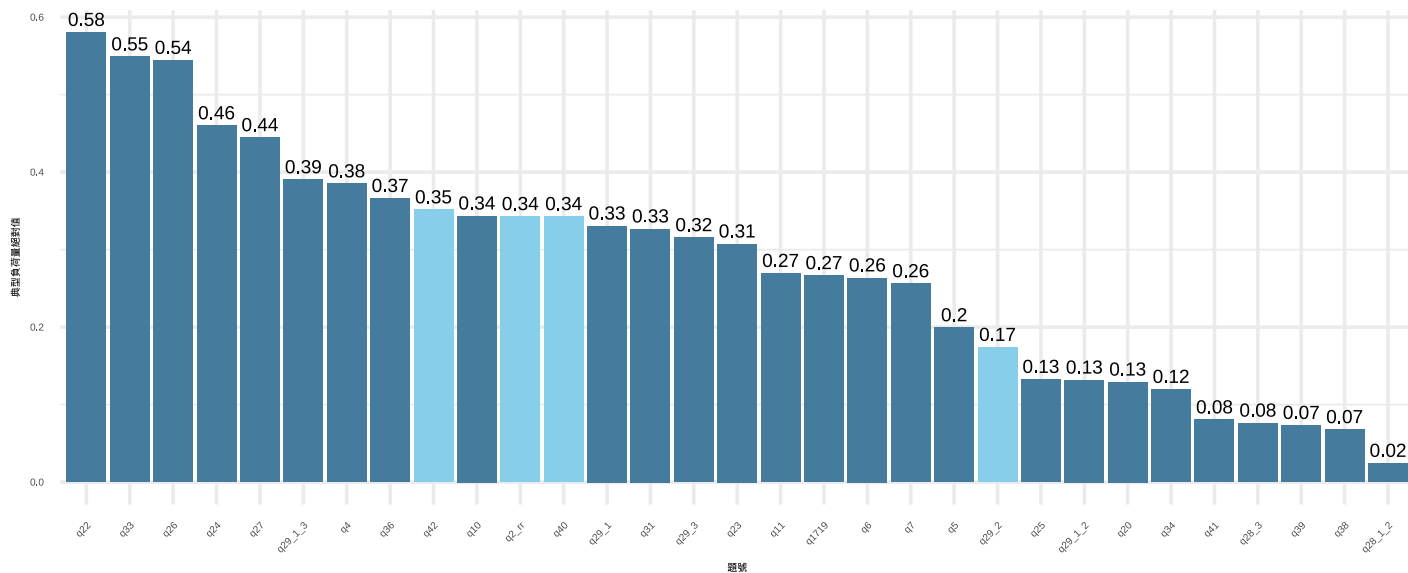
典型相關係數 0.5493481 0.3077433 0.2154761

[1] 0.7704 0.2891

[1] 0.3155 0.2306

[1] 0.2325 0.0274

[1] 0.0952 0.0218



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](#)