Task 2

EDA

By 白宸宇

Data

	原始資料		分類後資料	
	train_data_feeds.csv	train_data_ads.csv	potential_customers.csv	non_potential_customers.csv
欄位數	28 (u_userId)	35 (user_id)	28 (u_userId)	28 (u_userId)
資料筆數	3,227,732	7,675,517	2,475,843	751,889
Unique Value	180,123	65,297	65,297	114,826
說明	顧客看到廣告行為的資料	有點擊廣告的顧客資料 Ads中的所有資料都是潛在客戶 其中由65,297個客戶貢獻了7,675,517筆數據	feeds.csv中user_id同時出現在feeds與ads中的各比資料	feeds.csv中user_id僅出現在feeds中的各比資料

變數摘要

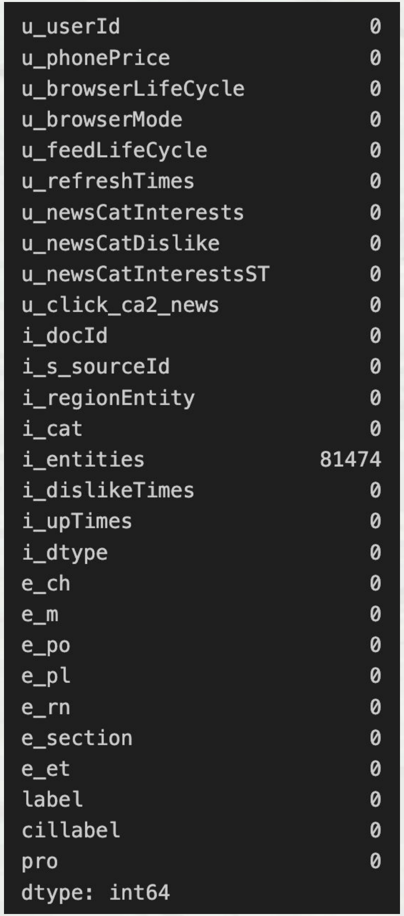
```
#   Column                 Dtype
---  ------                -----
 0   u_userId              int64
 1   u_phonePrice          int64
 2   u_browserLifeCycle    int64
 3   u_browserMode         int64
 4   u_feedLifeCycle       int64
 5   u_refreshTimes        int64
 6   u_newsCatInterests    object
 7   u_newsCatDislike      object
 8   u_newsCatInterestsST  object
 9   u_click_ca2_news      object
 10  i_docId               object
 11  i_s_sourceId          object
 12  i_regionEntity        int64
 13  i_cat                 int64
 14  i_entities            object
 15  i_dislikeTimes        int64
 16  i_upTimes             int64
 17  i_dtype               int64
 18  e_ch                  int64
 19  e_m                   int64
 20  e_po                  int64
 21  e_pl                  int64
 22  e_rn                  int64
 23  e_section             int64
 24  label                 int64
 25  cillabel              int64
 26  pro                   int64
```

經過觀察後發現object的欄位大
多屬於一串序號，無法進行分析

因此將object類別的欄位Drop掉

```
#   Column                 Dtype
---  ------                -----
 0   u_userId              int64
 1   u_phonePrice          int64
 2   u_browserLifeCycle    int64
 3   u_browserMode         int64
 4   u_feedLifeCycle       int64
 5   u_refreshTimes        int64
 6   i_regionEntity        int64
 7   i_cat                 int64
 8   i_dislikeTimes        int64
 9   i_upTimes             int64
 10  i_dtype               int64
 11  e_ch                  int64
 12  e_m                   int64
 13  e_po                  int64
 14  e_pl                  int64
 15  e_rn                  int64
 16  e_section             int64
 17  e_et                  int64
 18  label                 int64
 19  cillabel              int64
 20  pro                   int64
```

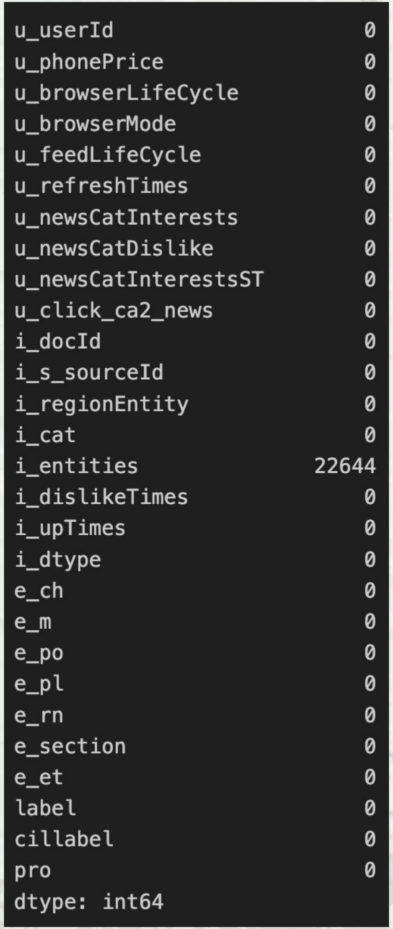
Potential Customers



```
u_userId                 0
u_phonePrice             0
u_browserLifeCycle       0
u_browserMode            0
u_feedLifeCycle          0
u_refreshTimes           0
u_newsCatInterests       0
u_newsCatDislike         0
u_newsCatInterestsST     0
u_click_ca2_news         0
i_docId                  0
i_s_sourceId             0
i_regionEntity           0
i_cat                    0
i_entities           81474
i_dislikeTimes           0
i_upTimes                0
i_dtype                  0
e_ch                     0
e_m                      0
e_po                     0
e_pl                     0
e_rn                     0
e_section                0
e_et                     0
label                    0
cillabel                 0
pro                      0
dtype: int64
```

觀察缺失值
僅i_entities有缺失值,
但由於屬於Object物件，故不做處理

Non-Potential Customers



```
u_userId                 0
u_phonePrice             0
u_browserLifeCycle       0
u_browserMode            0
u_feedLifeCycle          0
u_refreshTimes           0
u_newsCatInterests       0
u_newsCatDislike         0
u_newsCatInterestsST     0
u_click_ca2_news         0
i_docId                  0
i_s_sourceId             0
i_regionEntity           0
i_cat                    0
i_entities           22644
i_dislikeTimes           0
i_upTimes                0
i_dtype                  0
e_ch                     0
e_m                      0
e_po                     0
e_pl                     0
e_rn                     0
e_section                0
e_et                     0
label                    0
cillabel                 0
pro                      0
dtype: int64
```

Variable Definition

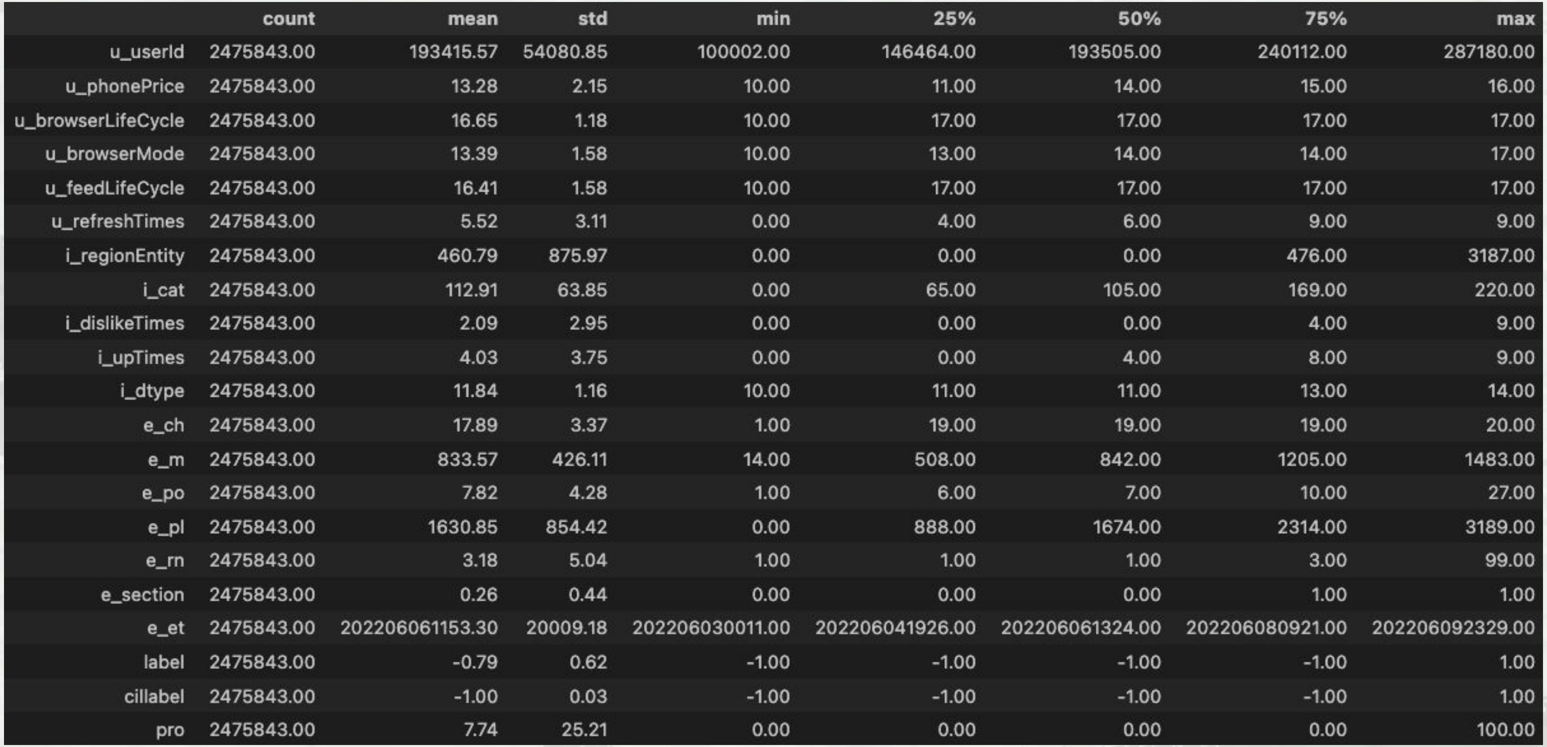
			潛在客戶	非潛在客戶
變數名稱	定義	類型	可能值	
u_userId	使用者的ID	離散	65297	114826
u_phonePrice	使用者裝置的價格	離散	7	7
u_browserLifeCycle	瀏覽器上的使用者參與度	離散	8	8
u_browserMode	遊覽器服務類型	離散	8	8
u_feedLifeCycle	用戶對新動態(feeds)的參與度	離散	8	8
u_refreshTimes	每天有效新動態(feeds)更新的平均數量。	離散	10	10
i_regionEntity	文章地域詞的ID	離散	374	364
i_cat	文章類別的ID	離散	208	207
i_dislikeTimes	文章負面反饋數	離散	10	10
i_upTimes	文章點讚數	離散	10	10

			潛在客戶	非潛在客戶
變數名稱	定義	類別	可能值	
i_dtype	文章展現形式	離散	5	5
e_ch	頻道	離散	19	19
e_m	事件來源設備機型	離散	256	262
e_po	第幾位	離散	27	27
e_pl	拜訪地	離散	3011	3089
e_rn	第幾刷	離散	99	99
e_section	訊息場景類型	離散	2	2
e_et	時間戳	離散	3653	3561
label	是否點擊	離散	2 (-1, 1)	2
cillabel	是否點讚	離散	2 (-1, 1)	2
pro	文章瀏覽進度	離散	82	35

1.分類後資料 EDA

基本統計量

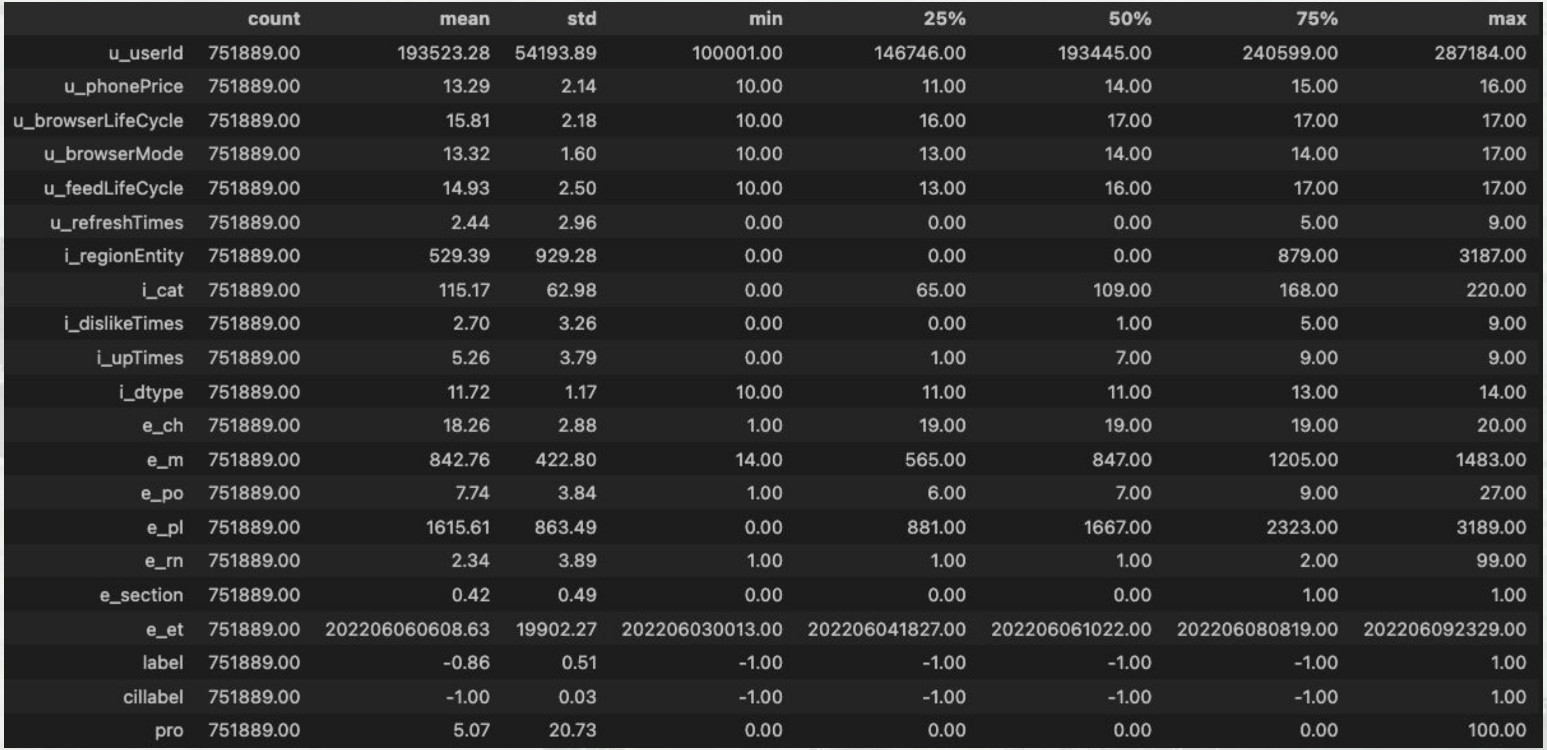
Potential Customers



	count	mean	std	min	25%	50%	75%	max
u_userId	2475843.00	193415.57	54080.85	100002.00	146464.00	193505.00	240112.00	287180.00
u_phonePrice	2475843.00	13.28	2.15	10.00	11.00	14.00	15.00	16.00
u_browserLifeCycle	2475843.00	16.65	1.18	10.00	17.00	17.00	17.00	17.00
u_browserMode	2475843.00	13.39	1.58	10.00	13.00	14.00	14.00	17.00
u_feedLifeCycle	2475843.00	16.41	1.58	10.00	17.00	17.00	17.00	17.00
u_refreshTimes	2475843.00	5.52	3.11	0.00	4.00	6.00	9.00	9.00
i_regionEntity	2475843.00	460.79	875.97	0.00	0.00	0.00	476.00	3187.00
i_cat	2475843.00	112.91	63.85	0.00	65.00	105.00	169.00	220.00
i_dislikeTimes	2475843.00	2.09	2.95	0.00	0.00	0.00	4.00	9.00
i_upTimes	2475843.00	4.03	3.75	0.00	0.00	4.00	8.00	9.00
i_dtype	2475843.00	11.84	1.16	10.00	11.00	11.00	13.00	14.00
e_ch	2475843.00	17.89	3.37	1.00	19.00	19.00	19.00	20.00
e_m	2475843.00	833.57	426.11	14.00	508.00	842.00	1205.00	1483.00
e_po	2475843.00	7.82	4.28	1.00	6.00	7.00	10.00	27.00
e_pl	2475843.00	1630.85	854.42	0.00	888.00	1674.00	2314.00	3189.00
e_rn	2475843.00	3.18	5.04	1.00	1.00	1.00	3.00	99.00
e_section	2475843.00	0.26	0.44	0.00	0.00	0.00	1.00	1.00
e_et	2475843.00	202206061153.30	20009.18	202206030011.00	202206041926.00	202206061324.00	202206080921.00	202206092329.00
label	2475843.00	-0.79	0.62	-1.00	-1.00	-1.00	-1.00	1.00
cillabel	2475843.00	-1.00	0.03	-1.00	-1.00	-1.00	-1.00	1.00
pro	2475843.00	7.74	25.21	0.00	0.00	0.00	0.00	100.00

基本統計量

Non-Potential Customers

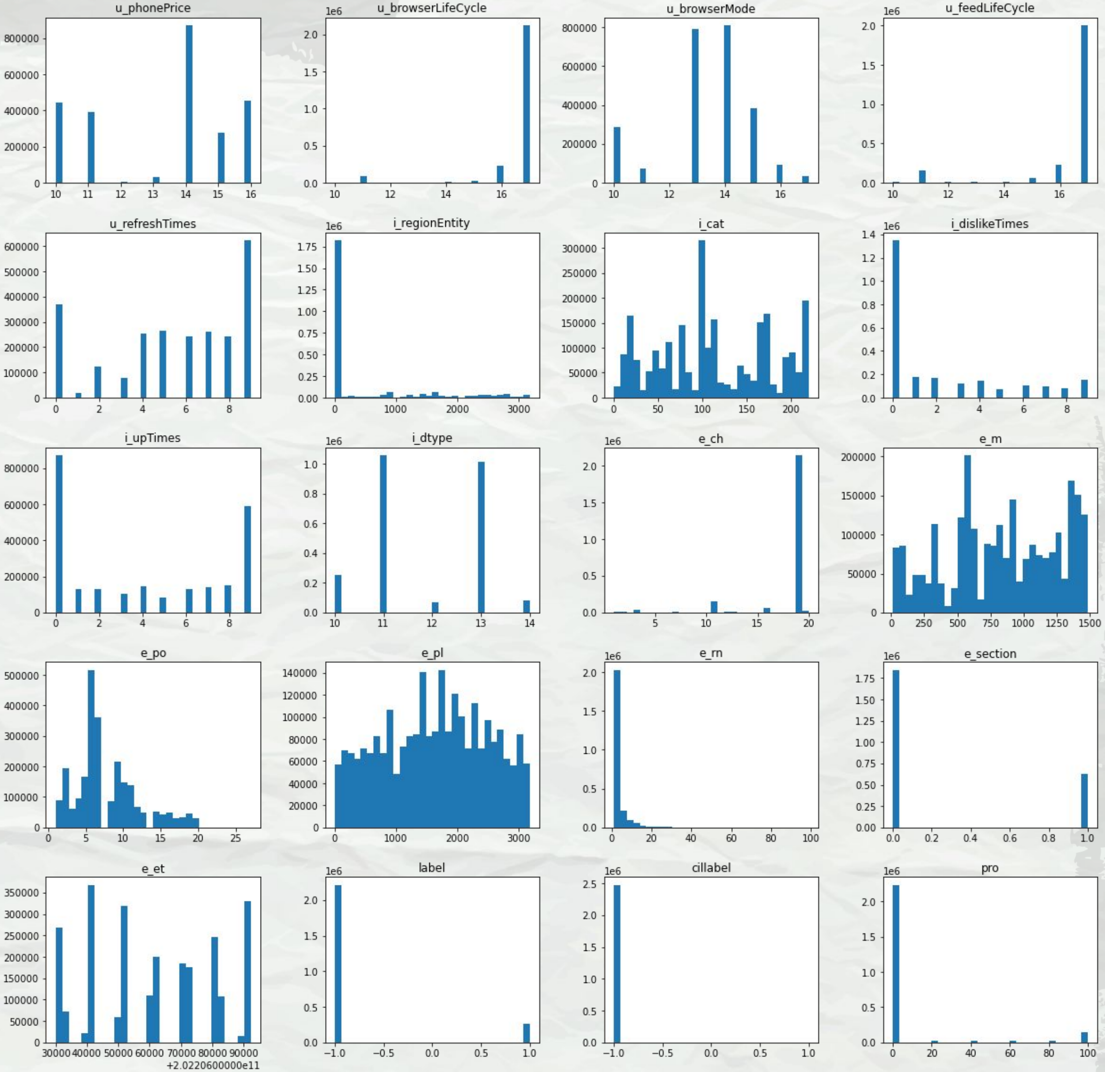


	count	mean	std	min	25%	50%	75%	max
u_userId	751889.00	193523.28	54193.89	100001.00	146746.00	193445.00	240599.00	287184.00
u_phonePrice	751889.00	13.29	2.14	10.00	11.00	14.00	15.00	16.00
u_browserLifeCycle	751889.00	15.81	2.18	10.00	16.00	17.00	17.00	17.00
u_browserMode	751889.00	13.32	1.60	10.00	13.00	14.00	14.00	17.00
u_feedLifeCycle	751889.00	14.93	2.50	10.00	13.00	16.00	17.00	17.00
u_refreshTimes	751889.00	2.44	2.96	0.00	0.00	0.00	5.00	9.00
i_regionEntity	751889.00	529.39	929.28	0.00	0.00	0.00	879.00	3187.00
i_cat	751889.00	115.17	62.98	0.00	65.00	109.00	168.00	220.00
i_dislikeTimes	751889.00	2.70	3.26	0.00	0.00	1.00	5.00	9.00
i_upTimes	751889.00	5.26	3.79	0.00	1.00	7.00	9.00	9.00
i_dtype	751889.00	11.72	1.17	10.00	11.00	11.00	13.00	14.00
e_ch	751889.00	18.26	2.88	1.00	19.00	19.00	19.00	20.00
e_m	751889.00	842.76	422.80	14.00	565.00	847.00	1205.00	1483.00
e_po	751889.00	7.74	3.84	1.00	6.00	7.00	9.00	27.00
e_pl	751889.00	1615.61	863.49	0.00	881.00	1667.00	2323.00	3189.00
e_rn	751889.00	2.34	3.89	1.00	1.00	1.00	2.00	99.00
e_section	751889.00	0.42	0.49	0.00	0.00	0.00	1.00	1.00
e_et	751889.00	202206060608.63	19902.27	202206030013.00	202206041827.00	202206061022.00	202206080819.00	202206092329.00
label	751889.00	-0.86	0.51	-1.00	-1.00	-1.00	-1.00	1.00
cillabel	751889.00	-1.00	0.03	-1.00	-1.00	-1.00	-1.00	1.00
pro	751889.00	5.07	20.73	0.00	0.00	0.00	0.00	100.00

直方圖

觀察類別型各欄位值的分佈

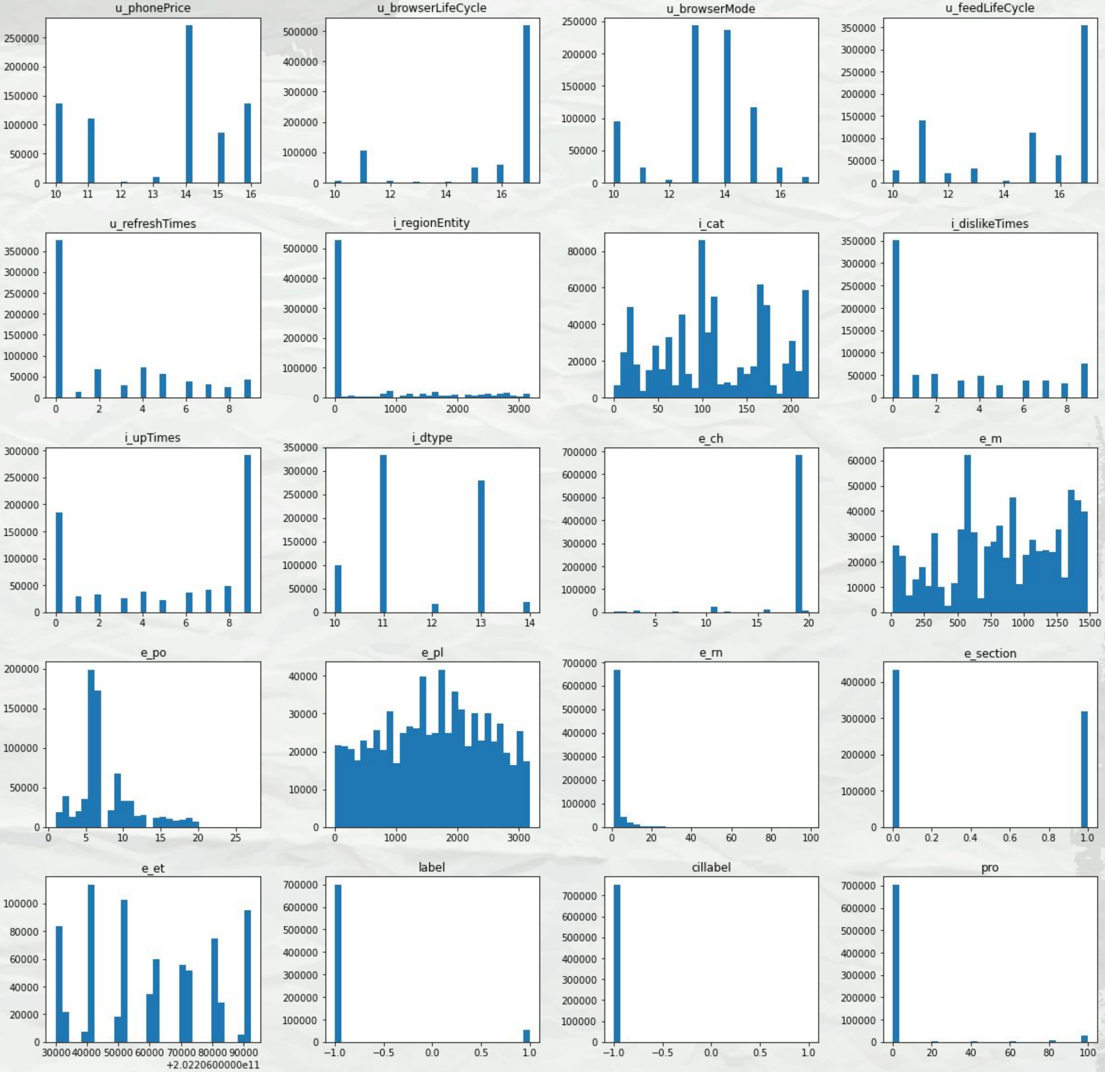
Potential Customers



直方圖

觀察類別型各欄位值的分佈

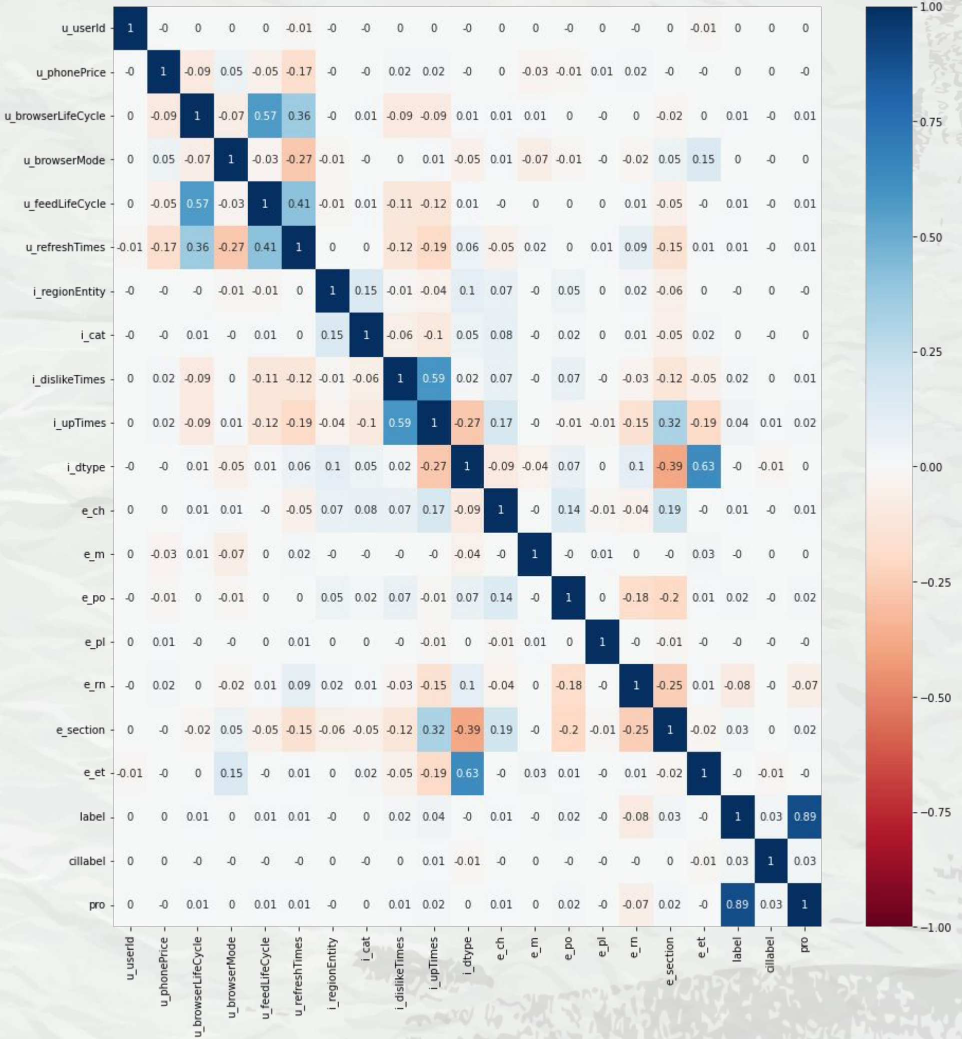
Non-
Potential Customers



相關係數熱點圖

觀察個欄位之間的相關係數

Potential Customers

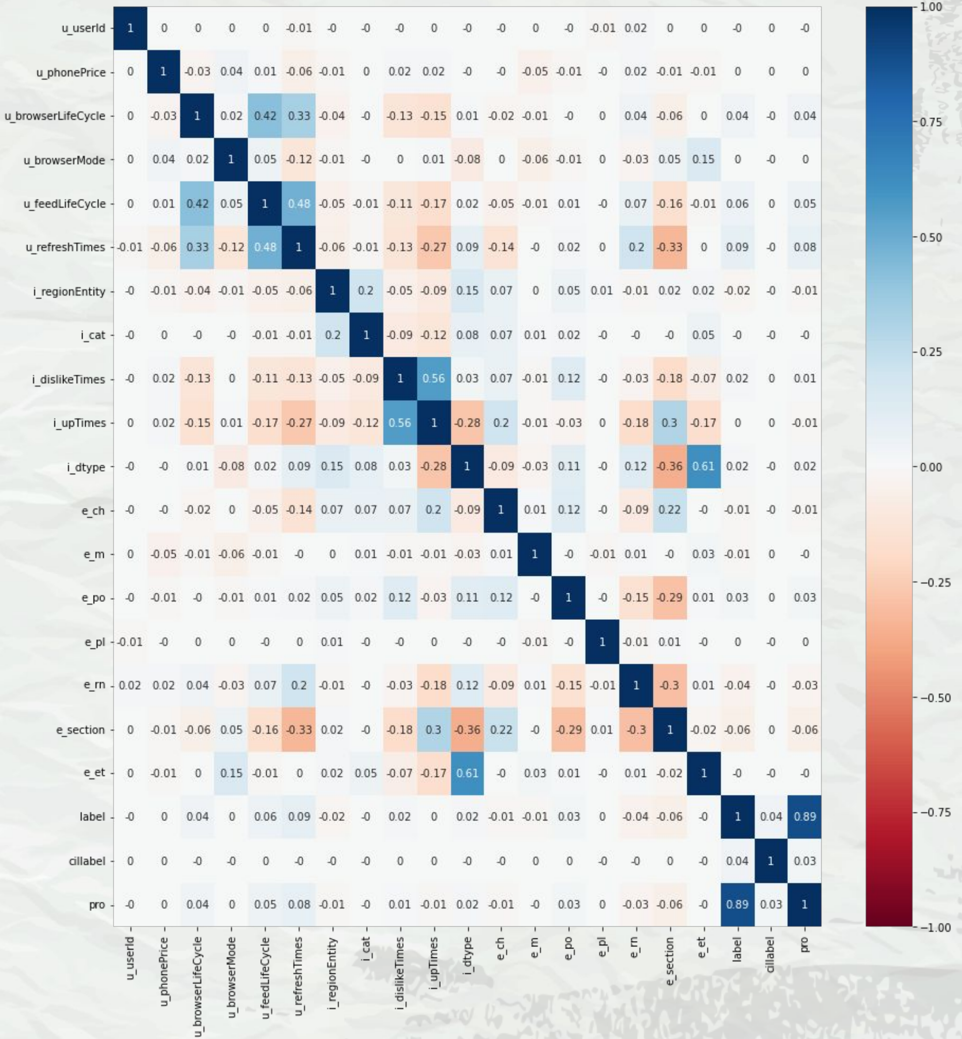


	u_userId	u_phonePrice	u_browserLifeCycle	u_browserMode	u_feedLifeCycle	u_refreshTimes	i_regionEntity	i_cat	i_dislikeTimes	i_upTimes	i_dtype	e_ch	e_m	e_po	e_pl	e_rn	e_section	e_et	label	cillabel	pro
u_userId	1	-0	0	0	0	-0.01	-0	-0	0	0	-0	0	0	-0	0	-0	0	-0.01	0	0	0
u_phonePrice	-0	1	-0.09	0.05	-0.05	-0.17	-0	-0	0.02	0.02	-0	0	-0.03	-0.01	0.01	0.02	-0	-0	0	0	-0
u_browserLifeCycle	0	-0.09	1	-0.07	0.57	0.36	-0	0.01	-0.09	-0.09	0.01	0.01	0.01	0	-0	0	-0.02	0	0.01	-0	0.01
u_browserMode	0	0.05	-0.07	1	-0.03	-0.27	-0.01	-0	0	0.01	-0.05	0.01	-0.07	-0.01	-0	-0.02	0.05	0.15	0	-0	0
u_feedLifeCycle	0	-0.05	0.57	-0.03	1	0.41	-0.01	0.01	-0.11	-0.12	0.01	-0	0	0	0	0.01	-0.05	-0	0.01	-0	0.01
u_refreshTimes	-0.01	-0.17	0.36	-0.27	0.41	1	0	0	-0.12	-0.19	0.06	-0.05	0.02	0	0.01	0.09	-0.15	0.01	0.01	-0	0.01
i_regionEntity	-0	-0	-0	-0.01	-0.01	0	1	0.15	-0.01	-0.04	0.1	0.07	-0	0.05	0	0.02	-0.06	0	-0	0	-0
i_cat	-0	-0	0.01	-0	0.01	0	0.15	1	-0.06	-0.1	0.05	0.08	-0	0.02	0	0.01	-0.05	0.02	0	-0	0
i_dislikeTimes	0	0.02	-0.09	0	-0.11	-0.12	-0.01	-0.06	1	0.59	0.02	0.07	-0	0.07	-0	-0.03	-0.12	-0.05	0.02	0	0.01
i_upTimes	0	0.02	-0.09	0.01	-0.12	-0.19	-0.04	-0.1	0.59	1	-0.27	0.17	-0	-0.01	-0.01	-0.15	0.32	-0.19	0.04	0.01	0.02
i_dtype	-0	-0	0.01	-0.05	0.01	0.06	0.1	0.05	0.02	-0.27	1	-0.09	-0.04	0.07	0	0.1	-0.39	0.63	-0	-0.01	0
e_ch	0	0	0.01	0.01	-0	-0.05	0.07	0.08	0.07	0.17	-0.09	1	-0	0.14	-0.01	-0.04	0.19	-0	0.01	-0	0.01
e_m	0	-0.03	0.01	-0.07	0	0.02	-0	-0	-0	-0	-0.04	-0	1	-0	0.01	0	-0	0.03	-0	0	0
e_po	-0	-0.01	0	-0.01	0	0	0.05	0.02	0.07	-0.01	0.07	0.14	-0	1	0	-0.18	-0.2	0.01	0.02	-0	0.02
e_pl	0	0.01	-0	-0	0	0.01	0	0	-0	-0.01	0	-0.01	0.01	0	1	-0	-0.01	-0	-0	-0	-0
e_rn	-0	0.02	0	-0.02	0.01	0.09	0.02	0.01	-0.03	-0.15	0.1	-0.04	0	-0.18	-0	1	-0.25	0.01	-0.08	-0	-0.07
e_section	0	-0	-0.02	0.05	-0.05	-0.15	-0.06	-0.05	-0.12	0.32	-0.39	0.19	-0	-0.2	-0.01	-0.25	1	-0.02	0.03	0	0.02
e_et	-0.01	-0	0	0.15	-0	0.01	0	0.02	-0.05	-0.19	0.63	-0	0.03	0.01	-0	0.01	-0.02	1	-0	-0.01	-0
label	0	0	0.01	0	0.01	0.01	-0	0	0.02	0.04	-0	0.01	-0	0.02	-0	-0.08	0.03	-0	1	0.03	0.89
cillabel	0	0	-0	-0	-0	-0	0	-0	0	0.01	-0.01	-0	0	-0	-0	-0	0	-0.01	0.03	1	0.03
pro	0	-0	0.01	0	0.01	0.01	-0	0	0.01	0.02	0	0.01	0	0.02	-0	-0.07	0.02	-0	0.89	0.03	1

相關係數熱點圖

觀察個欄位之間的相關係數

Non-
Potential Customers

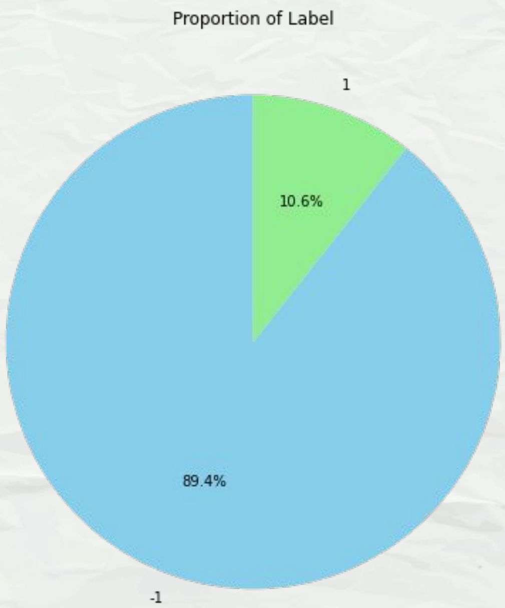


	u_userId	u_phonePrice	u_browserLifeCycle	u_browserMode	u_feedLifeCycle	u_refreshTimes	i_regionEntity	i_cat	i_dislikeTimes	i_upTimes	i_dtype	e_ch	e_m	e_po	e_pl	e_rn	e_section	e_et	label	cillabel	pro
u_userId	1	0	0	0	0	-0.01	-0	-0	-0	0	-0	-0	0	-0	-0.01	0.02	0	0	-0	0	-0
u_phonePrice	0	1	-0.03	0.04	0.01	-0.06	-0.01	0	0.02	0.02	-0	-0	-0.05	-0.01	-0	0.02	-0.01	-0.01	0	0	0
u_browserLifeCycle	0	-0.03	1	0.02	0.42	0.33	-0.04	-0	-0.13	-0.15	0.01	-0.02	-0.01	-0	0	0.04	-0.06	0	0.04	-0	0.04
u_browserMode	0	0.04	0.02	1	0.05	-0.12	-0.01	-0	0	0.01	-0.08	0	-0.06	-0.01	0	-0.03	0.05	0.15	0	-0	0
u_feedLifeCycle	0	0.01	0.42	0.05	1	0.48	-0.05	-0.01	-0.11	-0.17	0.02	-0.05	-0.01	0.01	-0	0.07	-0.16	-0.01	0.06	0	0.05
u_refreshTimes	-0.01	-0.06	0.33	-0.12	0.48	1	-0.06	-0.01	-0.13	-0.27	0.09	-0.14	-0	0.02	0	0.2	-0.33	0	0.09	-0	0.08
i_regionEntity	-0	-0.01	-0.04	-0.01	-0.05	-0.06	1	0.2	-0.05	-0.09	0.15	0.07	0	0.05	0.01	-0.01	0.02	0.02	-0.02	-0	-0.01
i_cat	-0	0	-0	-0	-0.01	-0.01	0.2	1	-0.09	-0.12	0.08	0.07	0.01	0.02	-0	-0	-0	0.05	-0	-0	-0
i_dislikeTimes	-0	0.02	-0.13	0	-0.11	-0.13	-0.05	-0.09	1	0.56	0.03	0.07	-0.01	0.12	-0	-0.03	-0.18	-0.07	0.02	0	0.01
i_upTimes	0	0.02	-0.15	0.01	-0.17	-0.27	-0.09	-0.12	0.56	1	-0.28	0.2	-0.01	-0.03	0	-0.18	0.3	-0.17	0	0	-0.01
i_dtype	-0	-0	0.01	-0.08	0.02	0.09	0.15	0.08	0.03	-0.28	1	-0.09	-0.03	0.11	-0	0.12	-0.36	0.61	0.02	-0	0.02
e_ch	-0	-0	-0.02	0	-0.05	-0.14	0.07	0.07	0.07	0.2	-0.09	1	0.01	0.12	-0	-0.09	0.22	-0	-0.01	-0	-0.01
e_m	0	-0.05	-0.01	-0.06	-0.01	-0	0	0.01	-0.01	-0.01	-0.03	0.01	1	-0	-0.01	0.01	-0	0.03	-0.01	0	-0
e_po	-0	-0.01	-0	-0.01	0.01	0.02	0.05	0.02	0.12	-0.03	0.11	0.12	-0	1	-0	-0.15	-0.29	0.01	0.03	0	0.03
e_pl	-0.01	-0	0	0	-0	0	0.01	-0	-0	0	-0	-0	-0.01	-0	1	-0.01	0.01	-0	0	-0	0
e_rn	0.02	0.02	0.04	-0.03	0.07	0.2	-0.01	-0	-0.03	-0.18	0.12	-0.09	0.01	-0.15	-0.01	1	-0.3	0.01	-0.04	-0	-0.03
e_section	0	-0.01	-0.06	0.05	-0.16	-0.33	0.02	-0	-0.18	0.3	-0.36	0.22	-0	-0.29	0.01	-0.3	1	-0.02	-0.06	0	-0.06
e_et	0	-0.01	0	0.15	-0.01	0	0.02	0.05	-0.07	-0.17	0.61	-0	0.03	0.01	-0	0.01	-0.02	1	-0	-0	-0
label	-0	0	0.04	0	0.06	0.09	-0.02	-0	0.02	0	0.02	-0.01	-0.01	0.03	0	-0.04	-0.06	-0	1	0.04	0.89
cillabel	0	0	-0	-0	0	-0	-0	-0	0	0	-0	-0	0	0	-0	-0	0	-0	0.04	1	0.03
pro	-0	0	0.04	0	0.05	0.08	-0.01	-0	0.01	-0.01	0.02	-0.01	-0	0.03	0	-0.03	-0.06	-0	0.89	0.03	1

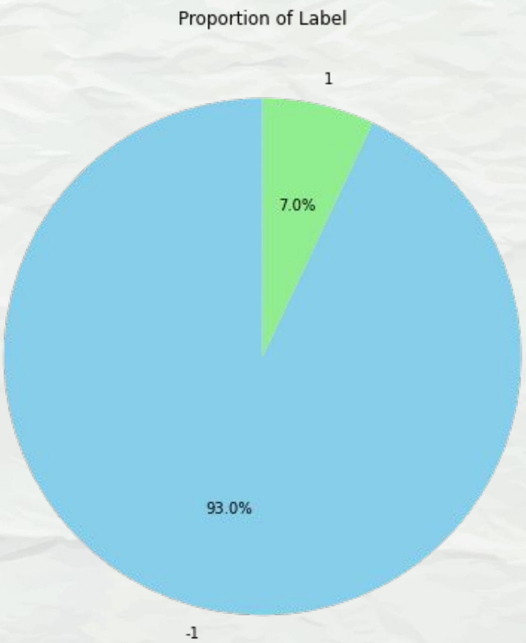
圓餅圖

觀察點擊的比率

潛在客戶

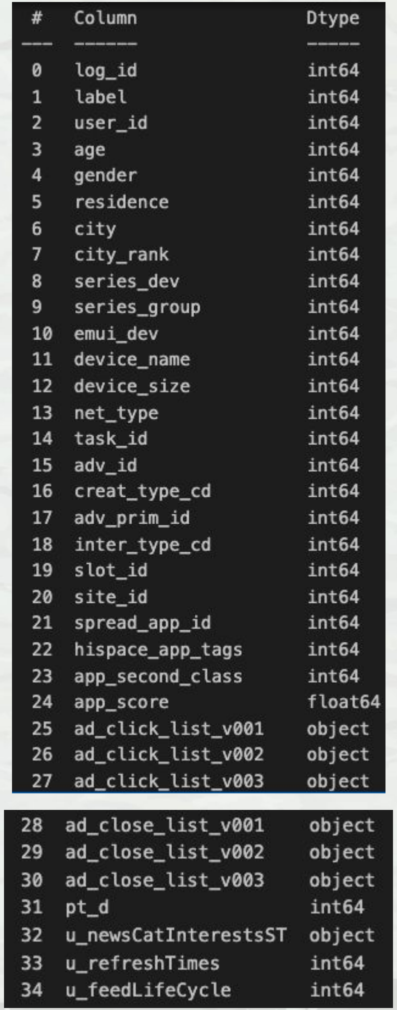


非潛在客戶



2.ads.csv之EDA

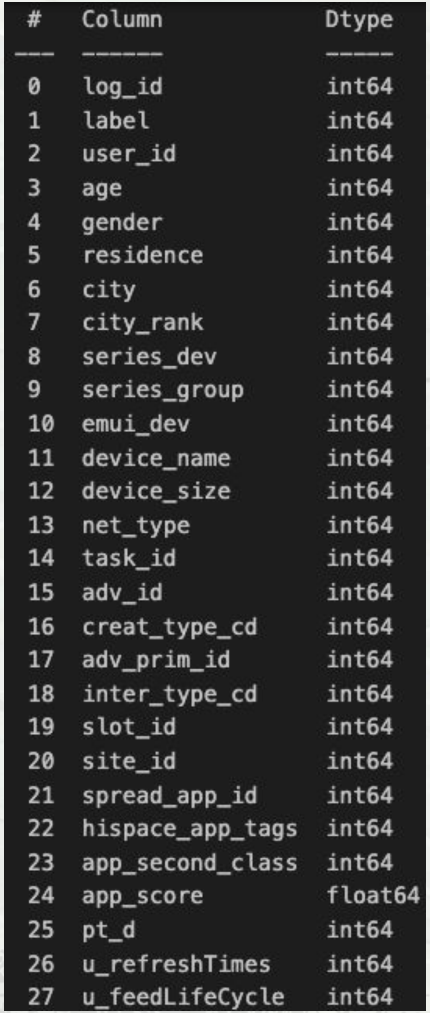
變數摘要



```
#   Column               Dtype
---  ------               -----
0   log_id               int64
1   label                int64
2   user_id              int64
3   age                  int64
4   gender               int64
5   residence            int64
6   city                 int64
7   city_rank            int64
8   series_dev           int64
9   series_group         int64
10  emui_dev             int64
11  device_name          int64
12  device_size          int64
13  net_type             int64
14  task_id              int64
15  adv_id               int64
16  creat_type_cd        int64
17  adv_prim_id          int64
18  inter_type_cd        int64
19  slot_id              int64
20  site_id              int64
21  spread_app_id        int64
22  hispace_app_tags     int64
23  app_second_class     int64
24  app_score            float64
25  ad_click_list_v001   object
26  ad_click_list_v002   object
27  ad_click_list_v003   object
28  ad_close_list_v001   object
29  ad_close_list_v002   object
30  ad_close_list_v003   object
31  pt_d                 int64
32  u_newsCatInterestsST object
33  u_refreshTimes       int64
34  u_feedLifeCycle      int64
```

經過觀察後發現object的欄位大多屬於一串序號，無法進行分析

因此將object類別的欄位Drop掉



```
#   Column            Dtype
---  ------            -----
0   log_id            int64
1   label             int64
2   user_id           int64
3   age               int64
4   gender            int64
5   residence         int64
6   city              int64
7   city_rank         int64
8   series_dev        int64
9   series_group      int64
10  emui_dev          int64
11  device_name       int64
12  device_size       int64
13  net_type          int64
14  task_id           int64
15  adv_id            int64
16  creat_type_cd     int64
17  adv_prim_id       int64
18  inter_type_cd     int64
19  slot_id           int64
20  site_id           int64
21  spread_app_id     int64
22  hispace_app_tags  int64
23  app_second_class  int64
24  app_score         float64
25  pt_d              int64
26  u_refreshTimes    int64
27  u_feedLifeCycle   int64
```

Variable Definition

變數名稱	定義	類型	可能值
label	使用者的ID	離散	2 (0,1)
user_id	使用者的ID	離散	65297
age	年紀	離散	8
gender	性別	離散	3
residence	永久居住地ID	離散	35
city	居住城市ID	離散	341
city_rank	居住城市等級	離散	4
series_dev	設備系列	離散	27
series_group	設備系列分組	離散	7
emui_dev	emui版本號碼	離散	27

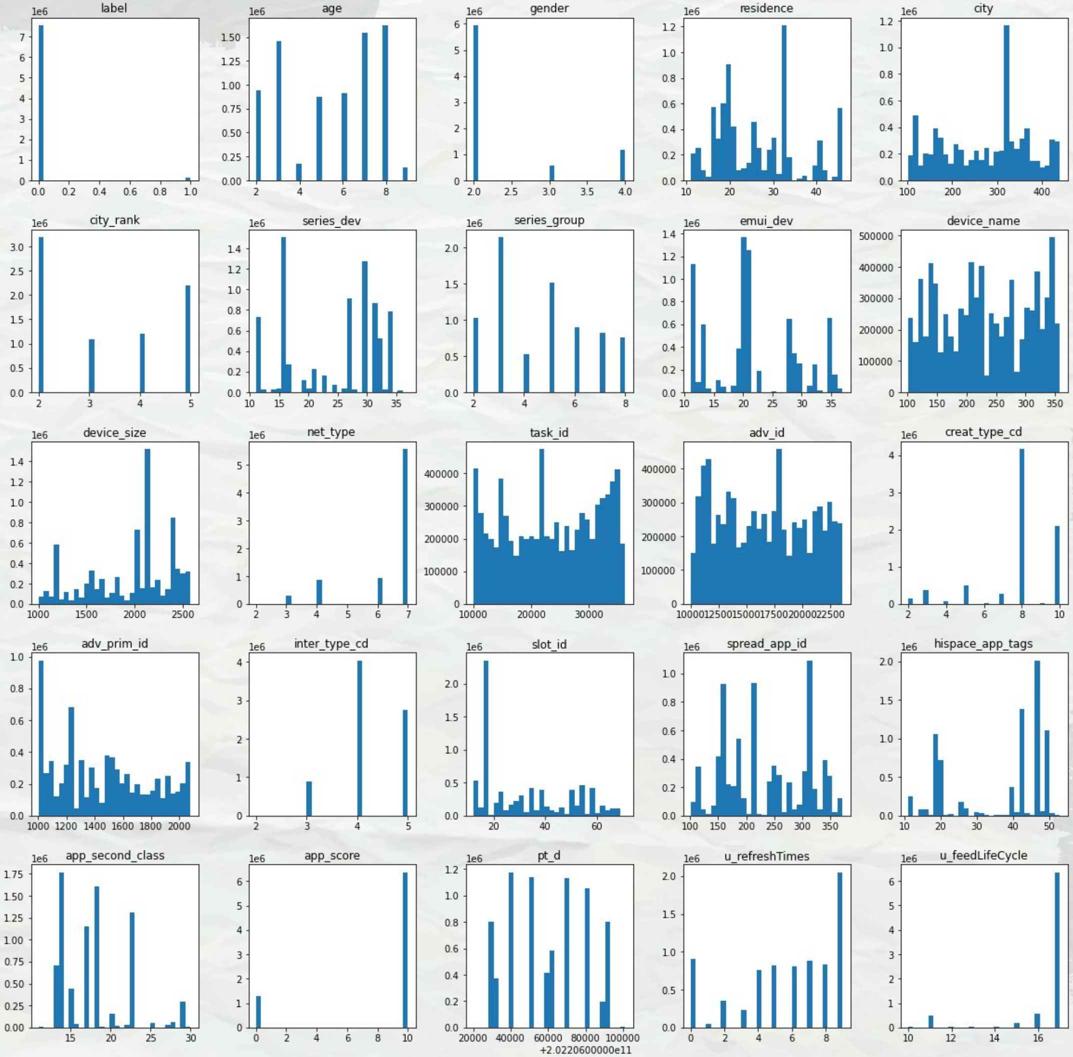
變數名稱	定義	類型	可能值
device_name	使用者手機機型	離散	256
device_size	使用者手機尺寸	離散	1547
net_type	行為時的網路狀態	離散	6
task_id	廣告任務ID	離散	11209
adv_id	廣告任務對應的素材ID	離散	12615
creat_type_cd	素材的創意類型ID	離散	9
adv_prim_id	廣告任務對應的廣告主ID	離散	545
inter_type_cd	廣告任務對應的素材的交互類型	離散	4
slot_id	廣告位ID	離散	60
site_id	媒體ID	離散	1
spread_app_id	投放廣告對應的App ID	離散	116

變數名稱	定義	類型	可能值
hispace_app_tags	廣告任務對應的App標籤	離散	43
app_second_class	廣告任務對應的App的二級分類	離散	20
app_score	App得分	離散	3
pt_d	時間戳	離散	5436
u_refreshTimes	每天有效新動態(feeds)更新的平均數量。	離散	10
u_feedLifeCycle	用戶對新動態(feeds)的參與度	離散	8

直方圖

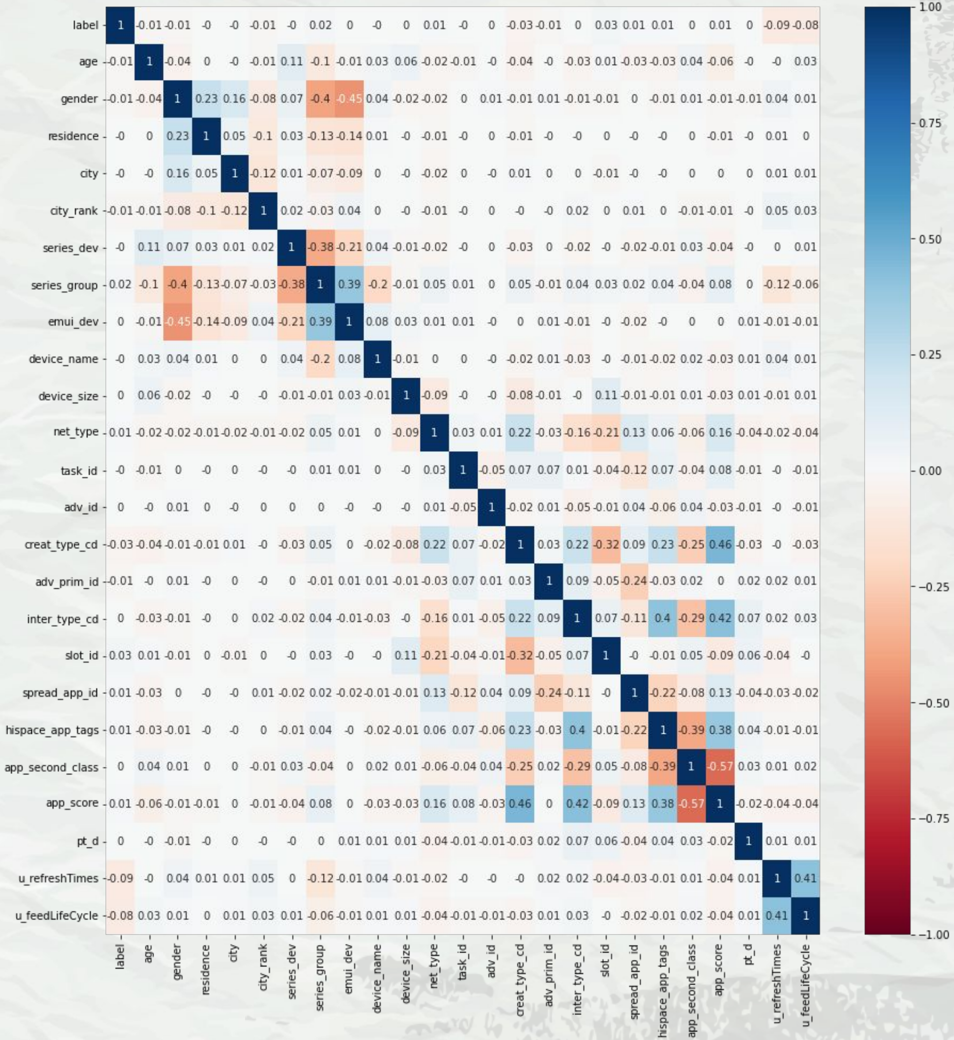
觀察類別型各欄位值的分佈

屬於Id的欄位觀察沒有意義



相關係數熱點圖

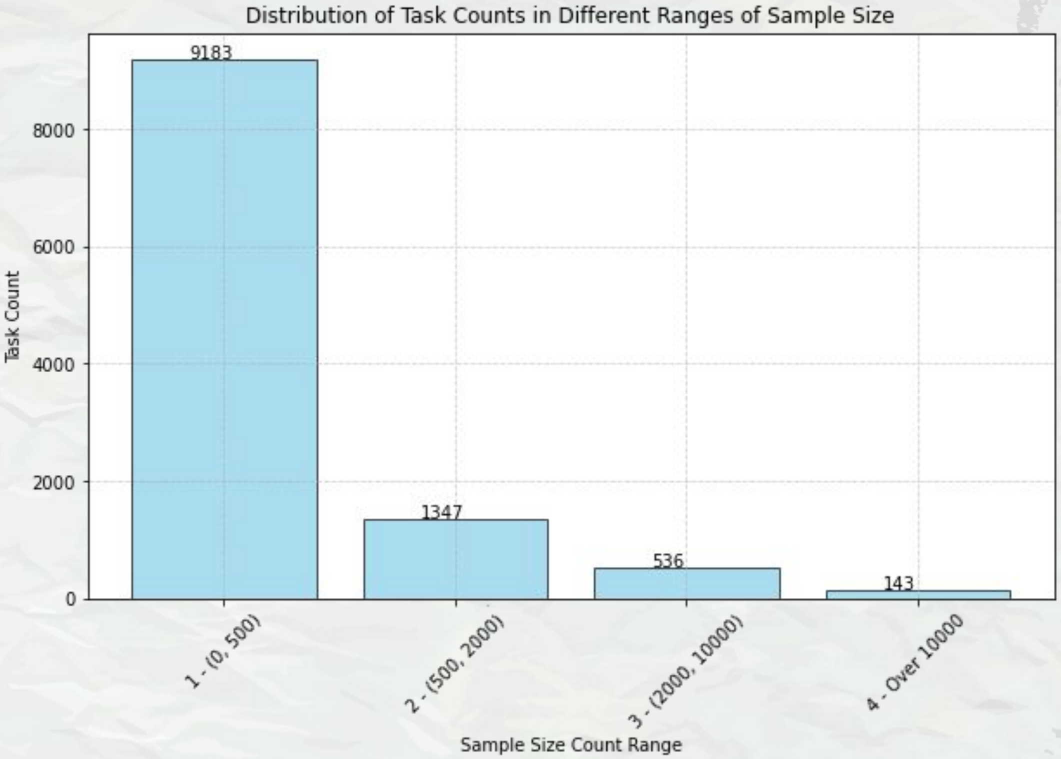
觀察個欄位之間的相關係數



	label	age	gender	residence	city	city_rank	series_dev	series_group	emui_dev	device_name	device_size	net_type	task_id	adv_id	creat_type_cd	adv_prim_id	inter_type_cd	slot_id	spread_app_id	hispace_app_tags	app_second_class	app_score	pt_d	u_refreshTimes	u_feedLifeCycle
label	1	-0.01	-0.01	-0	-0	-0.01	-0	0.02	0	-0	0	0.01	-0	0	-0.03	-0.01	0	0.03	0.01	0.01	0	0.01	0	-0.09	-0.08
age	-0.01	1	-0.04	0	-0	-0.01	0.11	-0.1	-0.01	0.03	0.06	-0.02	-0.01	-0	-0.04	-0	-0.03	0.01	-0.03	-0.03	0.04	-0.06	-0	-0	0.03
gender	-0.01	-0.04	1	0.23	0.16	-0.08	0.07	-0.4	-0.45	0.04	-0.02	-0.02	0	0.01	-0.01	0.01	-0.01	-0.01	0	-0.01	0.01	-0.01	-0.01	0.04	0.01
residence	-0	0	0.23	1	0.05	-0.1	0.03	-0.13	-0.14	0.01	-0	-0.01	-0	0	-0.01	-0	-0	0	-0	-0	0	-0.01	-0	0.01	0
city	-0	-0	0.16	0.05	1	-0.12	0.01	-0.07	-0.09	0	-0	-0.02	0	-0	0.01	0	0	-0.01	-0	-0	0	0	0	0.01	0.01
city_rank	-0.01	-0.01	-0.08	-0.1	-0.12	1	0.02	-0.03	0.04	0	-0	-0.01	-0	0	-0	-0	0.02	0	0.01	0	-0.01	-0.01	-0	0.05	0.03
series_dev	-0	0.11	0.07	0.03	0.01	0.02	1	-0.38	-0.21	0.04	-0.01	-0.02	-0	0	-0.03	0	-0.02	-0	-0.02	-0.01	0.03	-0.04	-0	0	0.01
series_group	0.02	-0.1	-0.4	-0.13	-0.07	-0.03	-0.38	1	0.39	-0.2	-0.01	0.05	0.01	0	0.05	-0.01	0.04	0.03	0.02	0.04	-0.04	0.08	0	-0.12	-0.06
emui_dev	0	-0.01	-0.45	-0.14	-0.09	0.04	-0.21	0.39	1	0.08	0.03	0.01	0.01	-0	0	0.01	-0.01	-0	-0.02	-0	0	0	0.01	-0.01	-0.01
device_name	-0	0.03	0.04	0.01	0	0	0.04	-0.2	0.08	1	-0.01	0	0	-0	-0.02	0.01	-0.03	-0	-0.01	-0.02	0.02	-0.03	0.01	0.04	0.01
device_size	0	0.06	-0.02	-0	-0	-0	-0.01	-0.01	0.03	-0.01	1	-0.09	-0	-0	-0.08	-0.01	-0	0.11	-0.01	-0.01	0.01	-0.03	0.01	-0.01	0.01
net_type	0.01	-0.02	-0.02	-0.01	-0.02	-0.01	-0.02	0.05	0.01	0	-0.09	1	0.03	0.01	0.22	-0.03	-0.16	-0.21	0.13	0.06	-0.06	0.16	-0.04	-0.02	-0.04
task_id	-0	-0.01	0	-0	0	-0	-0	0.01	0.01	0	-0	0.03	1	-0.05	0.07	0.07	0.01	-0.04	-0.12	0.07	-0.04	0.08	-0.01	-0	-0.01
adv_id	0	-0	0.01	0	-0	0	0	0	-0	-0	-0	0.01	-0.05	1	-0.02	0.01	-0.05	-0.01	0.04	-0.06	0.04	-0.03	-0.01	-0	-0.01
creat_type_cd	-0.03	-0.04	-0.01	-0.01	0.01	-0	-0.03	0.05	0	-0.02	-0.08	0.22	0.07	-0.02	1	0.03	0.22	-0.32	0.09	0.23	-0.25	0.46	-0.03	-0	-0.03
adv_prim_id	-0.01	-0	0.01	-0	0	-0	0	-0.01	0.01	0.01	-0.01	-0.03	0.07	0.01	0.03	1	0.09	-0.05	-0.24	-0.03	0.02	0	0.02	0.02	0.01
inter_type_cd	0	-0.03	-0.01	-0	0	0.02	-0.02	0.04	-0.01	-0.03	-0	-0.16	0.01	-0.05	0.22	0.09	1	0.07	-0.11	0.4	-0.29	0.42	0.07	0.02	0.03
slot_id	0.03	0.01	-0.01	0	-0.01	0	-0	0.03	-0	-0	0.11	-0.21	-0.04	-0.01	-0.32	-0.05	0.07	1	-0	-0.01	0.05	-0.09	0.06	-0.04	-0
spread_app_id	0.01	-0.03	0	-0	-0	0.01	-0.02	0.02	-0.02	-0.01	-0.01	0.13	-0.12	0.04	0.09	-0.24	-0.11	-0	1	-0.22	-0.08	0.13	-0.04	-0.03	-0.02
hispace_app_tags	0.01	-0.03	-0.01	-0	-0	0	-0.01	0.04	-0	-0.02	-0.01	0.06	0.07	-0.06	0.23	-0.03	0.4	-0.01	-0.22	1	-0.39	0.38	0.04	-0.01	-0.01
app_second_class	0	0.04	0.01	0	0	-0.01	0.03	-0.04	0	0.02	0.01	-0.06	-0.04	0.04	-0.25	0.02	-0.29	0.05	-0.08	-0.39	1	-0.57	0.03	0.01	0.02
app_score	0.01	-0.06	-0.01	-0.01	0	-0.01	-0.04	0.08	0	-0.03	-0.03	0.16	0.08	-0.03	0.46	0	0.42	-0.09	0.13	0.38	-0.57	1	-0.02	-0.04	-0.04
pt_d	0	-0	-0.01	-0	0	-0	-0	0	0.01	0.01	0.01	-0.04	-0.01	-0.01	-0.03	0.02	0.07	0.06	-0.04	0.04	0.03	-0.02	1	0.01	0.01
u_refreshTimes	-0.09	-0	0.04	0.01	0.01	0.05	0	-0.12	-0.01	0.04	-0.01	-0.02	-0	-0	-0	0.02	0.02	-0.04	-0.03	-0.01	0.01	-0.04	0.01	1	0.41
u_feedLifeCycle	-0.08	0.03	0.01	0	0.01	0.03	0.01	-0.06	-0.01	0.01	0.01	-0.04	-0.01	-0.01	-0.03	0.01	0.03	-0	-0.02	-0.01	0.02	-0.04	0.01	0.41	1

Task counts

觀察task_id中的數量分佈



Thanks for watching