Capital One

Is there a “best product” which works for all customers?

각 bkscore 군을 모두 만족시키는 “best product”는 찾을 수 없다. “Best”가 아니어도 각 bkscore의 고객을 상대적으로 평균 이상으로 만족시키는 offer를 찾을 수는 있다. 본 케이스에는 offer 1, 2, 4가 모든 고객을 평균 이상으로 만족시킬 수 있으나 개별 마케팅이 가능한 상관에서 굳이 평균 이상의 성과를 내는 offer를 사용할 필요는 없다. 비용을 줄여야 하거나 개인화 offer의 역량이 부족한 회사는 모든 고객에게 같은 offer를 보내 볼 수 있으나 가장 바람직한 방법은 아니다.

[Bkscore 150, 200, 250 and offers 1, 2, 4]

Bkscore 별 offer 1, 2, 4는 상위에 위치한 것을 볼 수 있다.

bkscore	offer	apr	fixvar	fee	people	sale	ltv	prob_pred	pred_profit
150	3	14.9	Fixed	0	4804	0	52	0.0392	2.0384
150	3	14.9	Fixed	0	196	1	52	0.0392	2.0384
150	4	14.9	Var.	0		0	62	0.029453	1.8261
150	4	14.9	Var.	0		1	62	0.029453	1.8261
150	11	19.8	Fixed	0	4928	0	100	0.0144	1.44
150	11	19.8	Fixed	0	72	1	100	0.0144	1.44
150	1	14.9	Fixed	20	4921	0	83	0.0158	1.3114
150	1	14.9	Fixed	20	79	1	83	0.0158	1.3114
150	12	19.8	Var.	0		0	110	0.010751	1.182562
150	12	19.8	Var.	0		1	110	0.010751	1.182562
150	2	14.9	Var.	20	4941	0	93	0.0118	1.0974
150	2	14.9	Var.	20	59	1	93	0.0118	1.0974
150	7	16.8	Fixed	0	4926	0	72	0.0148	1.0656
150	7	16.8	Fixed	0	74	1	72	0.0148	1.0656
150	8	16.8	Var.	0		0	82	0.01105	0.906126
150	8	16.8	Var.	0		1	82	0.01105	0.906126
150	9	19.8	Fixed	20		0	131	0.005716	0.748797
150	9	19.8	Fixed	20		1	131	0.005716	0.748797
150	5	16.8	Fixed	20		0	103	0.005876	0.605251
150	5	16.8	Fixed	20		1	103	0.005876	0.605251
150	10	19.8	Var.	20		0	141	0.004258	0.60036
150	10	19.8	Var.	20		1	141	0.004258	0.60036
150	6	16.8	Var.	20		0	113	0.004377	0.494646
150	6	16.8	Var.	20		1	113	0.004377	0.494646

bkscore	offer	apr	fixvar	fee	people	sale	ltv	prob_pred	pred_profit
200	4	14.9	Var.	0	4791	0	42	0.0418	1.7556
200	4	14.9	Var.	0	209	1	42	0.0418	1.7556
200	2	14.9	Var.	20		0	73	0.021456	1.566295
200	2	14.9	Var.	20		1	73	0.021456	1.566295
200	1	14.9	Fixed	20	4878	0	63	0.0244	1.5372
200	1	14.9	Fixed	20	122	1	63	0.0244	1.5372
200	3	14.9	Fixed	0	4763	0	32	0.0474	1.5168
200	3	14.9	Fixed	0	237	1	32	0.0474	1.5168
200	8	16.8	Var.	0		0	62	0.018635	1.155358
200	8	16.8	Var.	0		1	62	0.018635	1.155358
200	7	16.8	Fixed	0	4894	0	52	0.0212	1.1024
200	7	16.8	Fixed	0	106	1	52	0.0212	1.1024
200	11	19.8	Fixed	0	4936	0	80	0.0128	1.024
200	11	19.8	Fixed	0	64	1	80	0.0128	1.024
200	12	19.8	Var.	0		0	90	0.01124	1.011558
200	12	19.8	Var.	0		1	90	0.01124	1.011558
200	5	16.8	Fixed	20		0	83	0.010769	0.893857
200	5	16.8	Fixed	20		1	83	0.010769	0.893857
200	6	16.8	Var.	20		0	93	0.009454	0.879229
200	6	16.8	Var.	20		1	93	0.009454	0.879229
200	9	19.8	Fixed	20		0	111	0.006475	0.718715
200	9	19.8	Fixed	20		1	111	0.006475	0.718715
200	10	19.8	Var.	20		0	121	0.005681	0.687414
200	10	19.8	Var.	20		1	121	0.005681	0.687414

bkscore	offer	apr	fixvar	fee	people	sale	ltv	prob_pred	pred_profi
250	11	19.8	Fixed	0	4917	0	50	0.017417	0.870851
250	11	19.8	Fixed	0	83	1	50	0.017417	0.870851
250	12	19.8	Var.	0	4928	0	60	0.013583	0.814978
250	12	19.8	Var.	0	72	1	60	0.013583	0.814978
250	2	14.9	Var.	20	4912	0	43	0.018417	0.791932
250	2	14.9	Var.	20	88	1	43	0.018417	0.791932
250	1	14.9	Fixed	20	4878	0	33	0.023583	0.778238
250	1	14.9	Fixed	20	122	1	33	0.023583	0.778238
250	4	14.9	Var.	0		0	12	0.062075	0.744903
250	4	14.9	Var.	0		1	12	0.062075	0.744903
250	8	16.8	Var.	0	4886	0	32	0.0228	0.7296
250	8	16.8	Var.	0	114	1	32	0.0228	0.7296
250	7	16.8	Fixed	0		0	22	0.029159	0.641494
250	7	16.8	Fixed	0		1	22	0.029159	0.641494
250	5	16.8	Fixed	20		0	53	0.008443	0.447464
250	5	16.8	Fixed	20		1	53	0.008443	0.447464
250	6	16.8	Var.	20		0	63	0.006571	0.413972
250	6	16.8	Var.	20		1	63	0.006571	0.413972
250	9	19.8	Fixed	20	4975	0	81	0.005	0.405
250	9	19.8	Fixed	20	25	1	81	0.005	0.405
250	10	19.8	Var.	20		0	91	0.003889	0.353855
250	10	19.8	Var.	20		1	91	0.003889	0.353855
250	3	14.9	Fixed	0		0	2	0.078508	0.157016
250	3	14.9	Fixed	0		1	2	0.078508	0.157016

Describe and justify your testing strategy.

Test design

exhibit_all 파일을 보면 people 열의 일부 데이터가 수집되지 않은 것을 확인할 수 있었다. 이는 logistic regression 수행 시 frequency weight=people으로 보완해줄 예정이다.

[bkscore 190]은 apr 19.8에 대한 데이터가 없고 [bkscore 210]은 apr 14.9에 대한 데이터가 없기 때문에 bkscore 190과 210을 하나의 그룹으로 보고 분석을 진행하기로 하였다. 따라서 logistic regression을 bkscore이 [190, 210]인 그룹 하나 그리고 [bkscore 255]인 그룹 하나 총 2번 수행하였다. exhibt_all 과거 데이터의 전체 variable을 predictor로 놓고 logistic regression 분석을 하면 아래와 같다.

```
. logistic sale i.bkscore i.apr i.fixvar i.fee [fw=people]

Logistic regression                             Number of obs = 1,520,000
                                                LR chi2(6)    =   9185.88
                                                Prob > chi2   =    0.0000
Log likelihood = -132938.55                     Pseudo R2     =    0.0334
```

sale	Odds ratio	Std. err.	z	P>|z|	[95% conf.	interval]
bkscore						
210	1.232054	.0301803	8.52	0.000	1.174299	1.292649
255	1.424919	.0274412	18.39	0.000	1.372137	1.47973
apr						
16.8	.4712829	.0088915	-39.87	0.000	.4541742	.4890361
19.8	.2570848	.0060994	-57.25	0.000	.2454039	.2693216
fixvar						
Var.	.7407279	.0155073	-14.34	0.000	.7109494	.7717538
20.fee	.2896126	.0042501	-84.44	0.000	.2814012	.2980635
_cons	.0602449	.0009655	-175.30	0.000	.058382	.0621672

bkscore이 높아질수록, apr이 낮을수록, 고정금리일 때, 그리고 연회비가 없을 때 mailing을 통해 sale이 될 odds ratio가 증가하는 것을 확인할 수 있었다.

[bkscore 190, 210]에 대해 전체 variable로 logistic regression 수행 결과, apr과 fee variable이 유효한 interaction effect을 가지고 있는 것을 확인하였다.

```
Logistic regression                             Number of obs = 723,000
                                                LR chi2(7)    = 6636.97
                                                Prob > chi2   =  0.0000
Log likelihood = -59243.609                     Pseudo R2     =  0.0530
```

sale	Odds ratio	Std. err.	z	P>|z|	[95% conf.	interval]
210.bkscore	1.23265	.0714455	3.61	0.000	1.10028	1.380944
apr						
16.8	.5835585	.0375778	-8.36	0.000	.5143657	.6620592
19.8	.3297162	.0215858	-16.95	0.000	.2900106	.3748578
fixvar						
Var.	.6544598	.0293636	-9.45	0.000	.5993661	.7146177
apr#fixvar						
16.8#Var.	1	(empty)				
19.8#Var.	1	(empty)				
20.fee	.4257669	.0135267	-26.88	0.000	.4000636	.4531217
apr#fee						
16.8#20	.586858	.0264139	-11.84	0.000	.5373055	.6409805
19.8#20	.4644702	.0263106	-13.54	0.000	.4156619	.5190096
fixvar#fee						
Var.# 0	1	(empty)				
Var.#20	1	(omitted)				
_cons	.0515468	.0010677	-143.16	0.000	.0494961	.0536825

따라서 [bkscore 190, 210]에 대해서는 bkscore, fixvar, apr##fee interaction variable을 사용하여 분석하였다.

```
. logistic sale i.bkscore i.fixvar i.apr##i.fee if bkscore!=255 [fw=people]

Logistic regression                                   Number of obs = 723,000
                                                      LR chi2(7)    = 6636.97
                                                      Prob > chi2   =  0.0000
Log likelihood = -59243.609                           Pseudo R2     =  0.0530
```

sale	Odds ratio	Std. err.	z	P>|z|	[95% conf.	interval]
210.bkscore	1.23265	.0714455	3.61	0.000	1.10028	1.380944
fixvar						
Var.	.6544598	.0293636	-9.45	0.000	.5993661	.7146177
apr						
16.8	.5835585	.0375778	-8.36	0.000	.5143657	.6620592
19.8	.3297162	.0215858	-16.95	0.000	.2900106	.3748578
20.fee	.4257669	.0135267	-26.88	0.000	.4000636	.4531217
apr#fee						
16.8#20	.586858	.0264139	-11.84	0.000	.5373055	.6409805
19.8#20	.4644702	.0263106	-13.54	0.000	.4156619	.5190096
_cons	.0515468	.0010677	-143.16	0.000	.0494961	.0536825

[bkscore 255]에 대해 전체 variable로 logistic regression 수행 결과, 통계상으로 apr과 fee의 interaction variable이 유의한 결과를 나타내는 것으로 보였다.

```
Logistic regression                                   Number of obs = 797,000
                                                      LR chi2(5)    = 2736.09
                                                      Prob > chi2   =  0.0000
Log likelihood = -73563.835                           Pseudo R2     =  0.0183
```

sale	Odds ratio	Std. err.	z	P>|z|	[95% conf.	interval]
apr						
16.8	.3544532	.0293935	-12.51	0.000	.301281	.4170096
19.8	.2004388	.0154997	-20.78	0.000	.1722501	.2332407
fixvar						
Var.	.7175062	.0172152	-13.84	0.000	.6845461	.7520533
apr#fixvar						
16.8#Var.	1	(empty)				
19.8#Var.	1	(empty)				
20.fee	.2206913	.0174144	-19.15	0.000	.1890681	.2576038
apr#fee						
14.9# 0	1	(empty)				
16.8#20	1.331976	.1112754	3.43	0.001	1.130799	1.568943
19.8#20	1	(omitted)				
fee#fixvar						
0#Var.	1	(empty)				
20#Var.	1	(omitted)				
_cons	.1136012	.009133	-27.05	0.000	.09704	.1329888

하지만 결과를 해석해보면 이자율이 16.8과 연회비가 20% 때 offer를 수락하는 ratio 결과가 높게 나온 것이 일반적인 intuition와 일치하지 않는다. 이는 과거 1991년 11월 [bkscore 255] 고객에게 진행한 마케팅 프로모션 등의 외부 영향이 있었을 것으로 간주하고 interaction effect를 제외하여 logistic regression를 수행했다.

```
. logistic sale i.apr i.fee i.fixvar if bkscore==255 [fw=people]

Logistic regression                                     Number of obs = 797,000
                                                        LR chi2(4)    = 2723.54
                                                        Prob > chi2   =  0.0000
Log likelihood = -73570.107                             Pseudo R2     =  0.0182
```

sale	Odds ratio	Std. err.	z	P>|z|	[95% conf.	interval]
apr						
16.8	.4648981	.0110089	-32.34	0.000	.4438141	.4869837
19.8	.252447	.0088	-39.49	0.000	.2357754	.2702975
20.fee	.2841519	.0072662	-49.21	0.000	.2702614	.2987562
fixvar						
Var.	.7175062	.0172152	-13.84	0.000	.6845461	.7520533
_cons	.0882303	.0026332	-81.35	0.000	.0832174	.0935452

이 두 logistic regression을 통해 mailing offer에 반응할 확률을 예측하였다.

```
. predict prob_pred if bkscore!=255
(option pr assumed; Pr(sale))
(24 missing values generated)

. predict temp if bkscore==255
(option pr assumed; Pr(sale))
(48 missing values generated)

.
. replace prob_pred=temp if bkscore==255
(24 real changes made)

.
. drop temp
```

[bkscore 250]은 [bkscore 255]와 차이가 크지 않을 것으로 판단하고 [255]의 예측 값을 [250]에 적용하였다. [bkscore 200]은 [190, 210]의 평균 값을 계산하여 사용하다. [bkscore 150]은 [200, 250]의 차이가 50씩이어서 [200, 250]의 예측의 차이가 [150, 200]과 동일하게 날것으로 가정하여 [200, 250]의 예측 값 차이만큼 [200]에서 뺀 값을 [150]의 예측 값으로 적용하였다. 이는 엑셀로 계산하였고 결과는 아래와 같다.

offer	apr	fixvar	fee	LTV			Probability prediction			Expected LTV per customer		
				LTV_150	LTV_200	LTV_250	bk150_pred	bk200_pred	bk250_pred	bk150	bk200	bk250
1	14.9	Fixed	20	83	63	33	0.0233583	0.02390795	0.0244576	1.9387389	1.50620085	0.8071008
2	14.9	Var.	20	93	73	43	0.0138864	0.0157785	0.0176706	1.2914352	1.1518305	0.7598358
3	14.9	Fixed	0	52	32	2	0.0276862	0.05438155	0.0810769	1.4396824	1.7402096	0.1621538
4	14.9	Var.	0	62	42	12	0.0130213	0.03627905	0.0595368	0.8073206	1.5237201	0.7144416
5	16.8	Fixed	20	103	83	53	0.0051185	0.0083198	0.0115211	0.5272055	0.6905434	0.6106183
6	16.8	Var.	20	113	93	63	0.0026283	0.00546085	0.0082934	0.2969979	0.50785905	0.5224842
7	16.8	Fixed	0	72	52	22	0.0255534	0.03247765	0.0394019	1.8398448	1.6888378	0.8668418
8	16.8	Var.	0	82	62	32	0.0144089	0.0214991	0.0285893	1.1815298	1.3329442	0.9148576
9	19.8	Fixed	20	131	111	81	0.0011865	0.00373785	0.0062892	0.1554315	0.41490135	0.5094252
10	19.8	Var.	20	141	121	91	0.0003783	0.00244945	0.0045206	0.0533403	0.29638345	0.4113746
11	19.8	Fixed	0	100	80	50	0.0154436	0.0186159	0.0217882	1.54436	1.489272	1.08941
12	19.8	Var.	0	110	90	60	0.0087961	0.01226305	0.01573	0.967571	1.1036745	0.9438

Bkscore별 예상 LTV가 가장 높은 offer top 5을 추려보았을 때 아래와 같다.

[bkscore 150]

offer	apr	fixvar	fee	LTV_150	LTV_200	LTV_250	bk150
1	14.9	Fixed	20	83	63	33	1.9387389
7	16.8	Fixed	0	72	52	22	1.8398448
11	19.8	Fixed	0	100	80	50	1.54436
3	14.9	Fixed	0	52	32	2	1.4396824
2	14.9	Var.	20	93	73	43	1.2914352
8	16.8	Var.	0	82	62	32	1.1815298
12	19.8	Var.	0	110	90	60	0.967571
4	14.9	Var.	0	62	42	12	0.8073206
5	16.8	Fixed	20	103	83	53	0.5272055
6	16.8	Var.	20	113	93	63	0.2969979
9	19.8	Fixed	20	131	111	81	0.1554315
10	19.8	Var.	20	141	121	91	0.0533403

[bkscore 200]

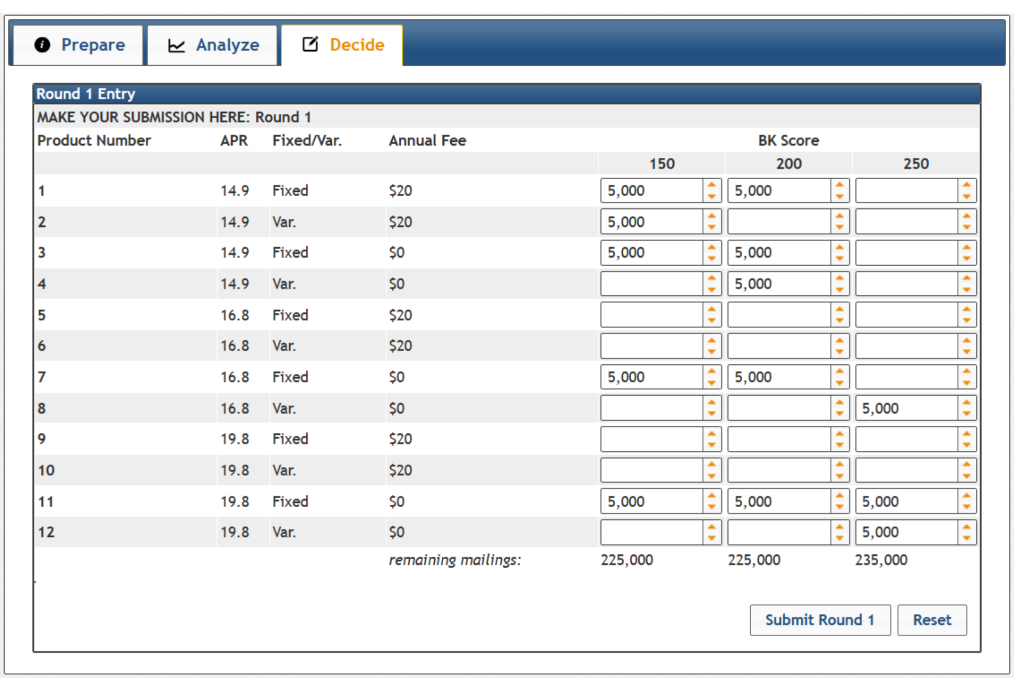
offer	apr	fixvar	fee	LTV_150	LTV_200	LTV_250	bk200
3	14.9	Fixed	0	52	32	2	1.7402096
7	16.8	Fixed	0	72	52	22	1.6888378
4	14.9	Var.	0	62	42	12	1.5237201
1	14.9	Fixed	20	83	63	33	1.50620085
11	19.8	Fixed	0	100	80	50	1.489272
8	16.8	Var.	0	82	62	32	1.3329442
2	14.9	Var.	20	93	73	43	1.1518305
12	19.8	Var.	0	110	90	60	1.1036745
5	16.8	Fixed	20	103	83	53	0.6905434
6	16.8	Var.	20	113	93	63	0.50785905
9	19.8	Fixed	20	131	111	81	0.41490135
10	19.8	Var.	20	141	121	91	0.29638345

[bkscore 250]

offer	apr	fixvar	fee	LTV_150	LTV_200	LTV_250	bk250
11	19.8	Fixed	0	100	80	50	1.08941
12	19.8	Var.	0	110	90	60	0.9438
8	16.8	Var.	0	82	62	32	0.9148576
7	16.8	Fixed	0	72	52	22	0.8668418
1	14.9	Fixed	20	83	63	33	0.8071008
2	14.9	Var.	20	93	73	43	0.7598358
4	14.9	Var.	0	62	42	12	0.7144416
5	16.8	Fixed	20	103	83	53	0.6106183
6	16.8	Var.	20	113	93	63	0.5224842
9	19.8	Fixed	20	131	111	81	0.5094252
10	19.8	Var.	20	141	121	91	0.4113746
3	14.9	Fixed	0	52	32	2	0.1621538

Round 1

위의 결과를 토대로 [150, 200]에는 Top 5 offer에 각각 5000개의 offer를 보냈고 LTV가 가장 낮은 [250]은 3개의 offer 보냈다. Round 2에서 알게 되었는데 variable apr, fixvar, fee에 대한 다양한 조합을 고려하여 offer를 보냈어야 했는데 그러지 않아 Round 1의 결과를 토대로 [250]의 logistic regression을 수행하였을 때 일부 apr과 fee에 대한 변수가 예측 값에 고려되지 못하는 상황이 발생하였다. 따라서 historical data의 offer 수락률을 참고하여 Round 2에 반영하였는데 이는 Round 2에서 상세히 설명하겠다.



Prepare | Analyze | Decide

Round 1 Entry

MAKE YOUR SUBMISSION HERE: Round 1

Product Number	APR	Fixed/Var.	Annual Fee	BK Score 150	BK Score 200	BK Score 250
1	14.9	Fixed	$20	5,000	5,000	
2	14.9	Var.	$20	5,000		
3	14.9	Fixed	$0	5,000	5,000	
4	14.9	Var.	$0		5,000	
5	16.8	Fixed	$20			
6	16.8	Var.	$20			
7	16.8	Fixed	$0	5,000	5,000	
8	16.8	Var.	$0			5,000
9	19.8	Fixed	$20			
10	19.8	Var.	$20			
11	19.8	Fixed	$0	5,000	5,000	5,000
12	19.8	Var.	$0			5,000
			remaining mailings:	225,000	225,000	235,000

Submit Round 1 | Reset

Round 1 결과

Prepare | Analyze | Decide

Economics | Round 1 Results

Round 1 Responses

	150		200		250	
	Sent	Responses	Sent	Responses	Sent	Responses
1	5,000	79	5,000	122	0	0
2	5,000	59	0	0	0	0
3	5,000	196	5,000	237	0	0
4	0	0	5,000	209	0	0
5	0	0	0	0	0	0
6	0	0	0	0	0	0
7	5,000	74	5,000	106	0	0
8	0	0	0	0	5,000	114
9	0	0	0	0	0	0
10	0	0	0	0	0	0
11	5,000	72	5,000	64	5,000	83
12	0	0	0	0	5,000	72

Round 1 Results

Solicitation Development	$17,000	Round 1 Number Sent	65,000
Mailing Costs	$800	Round 1 Number of Responses	1,487
Cost of Pieces Mailed	$32,500	Total Response Value	$81,562
Total Cost of Mailing	$50,300		
		Total Profit	$31,262

Round 1의 결과를 바탕으로 test_result 데이터에 people 칼럼을 업데이트하였고 각각의 bkscore에 대한 logistic regression을 수행하였다.

[bkscore 150]

```
. logistic sale i.apr i.fee i.fixvar if bkscore==150 [fw=people]

Logistic regression                                     Number of obs = 25,000
                                                        LR chi2(4)    = 114.22
                                                        Prob > chi2   = 0.0000
Log likelihood = -2315.6195                             Pseudo R2     = 0.0241

------------------------------------------------------------------------------
        sale | Odds ratio   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
         apr |
        16.8 |   .3682004   .0507886    -7.24   0.000     .2809776    .4824994
        19.8 |   .3581036   .0498807    -7.37   0.000     .2725486    .4705151
             |
      20.fee |   .3934782   .0530417    -6.92   0.000     .3021181    .5124654
             |
      fixvar |
        Var. |   .7438124     .12886    -1.71   0.088     .5296624    1.044546
       _cons |   .0407993   .0029731   -43.90   0.000     .0353692    .0470632
------------------------------------------------------------------------------
Note: _cons estimates baseline odds.

. predict prob_pred if bkscore==150
(option pr assumed; Pr(sale))
(48 missing values generated)
```

[bkscore 200]

```
. logistic sale i.apr i.fee i.fixvar if bkscore==200 [fw=people]

Logistic regression                                     Number of obs = 25,000
                                                        LR chi2(4)    = 150.59
                                                        Prob > chi2   = 0.0000
Log likelihood = -3251.4451                             Pseudo R2     = 0.0226

------------------------------------------------------------------------------
        sale | Odds ratio   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
         apr |
        16.8 |   .4352854    .051628    -7.01   0.000     .3449972    .5492028
        19.8 |   .2605776   .0370872    -9.45   0.000     .1971461    .3444181
             |
      20.fee |   .5026322   .0569353    -6.07   0.000     .4025603    .6275807
             |
      fixvar |
        Var. |   .8767027   .0851024    -1.36   0.175      .724812    1.060423
       _cons |   .0497586   .0033116   -45.09   0.000     .0436734    .0566915
------------------------------------------------------------------------------

. predict temp if bkscore==200
(option pr assumed; Pr(sale))
(48 missing values generated)

. replace prob_pred=temp if bkscore==200
(24 real changes made)

. drop temp
```

[bkscore 250]

위에 언급한 것과 같이 [250]의 partial design에 3개의 offer 밖에 넣지 않아 apr 14.9, fee 변수가 제외되고 분석되었다.

```
. logistic sale i.apr i.fee i.fixvar if bkscore==250 [fw=people]
note: 0.fee omitted because of collinearity.

Logistic regression                                     Number of obs = 15,000
                                                        LR chi2(2)    =  10.51
                                                        Prob > chi2   = 0.0052
Log likelihood = -1342.9914                             Pseudo R2     = 0.0039

------------------------------------------------------------------------------
        sale | Odds ratio   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
         apr |
        19.8 |   .6261961   .0951085    -3.08   0.002     .4649727    .8433218
       0.fee |          1  (omitted)
             |
      fixvar |
        Var. |   .8655336   .1404817    -0.89   0.374     .6296924    1.189705
       _cons |   .0269567   .0050661   -19.23   0.000     .0186507    .0389618
------------------------------------------------------------------------------
```

이를 보완하기 위해서 historical data를 다시 참고하여 offer 1, 2, 9의 반응률을 5000 : x의 비율로 변환하여 test_result에 추가해주었다. 이는 test 시점의 차이로 오류가 존재할 수 있으나 bkscore 250의 예측률을 높이기 위한 최선의 선택이었다.

[Historical data – 1, 2, 9 offer 수치 참고]

bkscore	offer	apr	fixvar	fee	people	sale	bk150	bk200	bk250
255	1	14.9	Fixed	20	172671	0	83	63	33
255	1	14.9	Fixed	20	4329	1	83	63	33
255	2	14.9	Var.	20	166996	0	93	73	43
255	2	14.9	Var.	20	3004	1	93	73	43
255	3	14.9	Fixed	0		0	52	32	2
255	3	14.9	Fixed	0		1	52	32	2
255	4	14.9	Var.	0		0	62	42	12
255	4	14.9	Var.	0		1	62	42	12
255	5	16.8	Fixed	20	252017	0	103	83	53
255	5	16.8	Fixed	20	2983	1	103	83	53
255	6	16.8	Var.	20		0	113	93	63
255	6	16.8	Var.	20		1	113	93	63
255	7	16.8	Fixed	0	62484	0	72	52	22
255	7	16.8	Fixed	0	2516	1	72	52	22
255	8	16.8	Var.	0		0	82	62	32
255	8	16.8	Var.	0		1	82	62	32
255	9	19.8	Fixed	20	34825	0	131	111	81
255	9	19.8	Fixed	20	175	1	131	111	81
255	10	19.8	Var.	20		0	141	121	91
255	10	19.8	Var.	20		1	141	121	91
255	11	19.8	Fixed	0	92885	0	100	80	50
255	11	19.8	Fixed	0	2115	1	100	80	50
255	12	19.8	Var.	0		0	110	90	60
255	12	19.8	Var.	0		1	110	90	60

[Test_result 수정 후 – 주황색 historical data에서 참고한 비율]

bkscore	offer	apr	fixvar	fee	people	sale	ltv	prob_pred
250	1	14.9	Fixed	20	4878	0	33	0.023583
250	1	14.9	Fixed	20	122	1	33	0.023583
250	2	14.9	Var.	20	4912	0	43	0.018417
250	2	14.9	Var.	20	88	1	43	0.018417
250	3	14.9	Fixed	0		0	2	0.078508
250	3	14.9	Fixed	0		1	2	0.078508
250	4	14.9	Var.	0		0	12	0.062075
250	4	14.9	Var.	0		1	12	0.062075
250	5	16.8	Fixed	20		0	53	0.008443
250	5	16.8	Fixed	20		1	53	0.008443
250	6	16.8	Var.	20		0	63	0.006571
250	6	16.8	Var.	20		1	63	0.006571
250	7	16.8	Fixed	0		0	22	0.029159
250	7	16.8	Fixed	0		1	22	0.029159
250	8	16.8	Var.	0	4886	0	32	0.0228
250	8	16.8	Var.	0	114	1	32	0.0228
250	9	19.8	Fixed	20	4975	0	81	0.005
250	9	19.8	Fixed	20	25	1	81	0.005
250	10	19.8	Var.	20		0	91	0.003889
250	10	19.8	Var.	20		1	91	0.003889
250	11	19.8	Fixed	0	4917	0	50	0.017417
250	11	19.8	Fixed	0	83	1	50	0.017417
250	12	19.8	Var.	0	4928	0	60	0.013583
250	12	19.8	Var.	0	72	1	60	0.013583

이 데이터로 [250]에 대한 logistic regression을 다시 실행하였다. Predict로 offer 반응률을 예측하였고 이를 LTV 값과 곱해주어 고객 당 예상되는 LTV를 계산하였다.

```
. logistic sale i.apr i.fee i.fixvar if bkscore==250 [fw=people]

Logistic regression                                     Number of obs = 30,000
                                                        LR chi2(4)    =  84.54
                                                        Prob > chi2   = 0.0000
Log likelihood = -2517.008                              Pseudo R2     = 0.0165

------------------------------------------------------------------------------
        sale | Odds ratio   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
         apr |
       16.8  |    .352534    .086901    -4.23   0.000     .2174585    .5715122
       19.8  |   .2080577   .0452301    -7.22   0.000     .1358749    .3185873
             |
      20.fee |   .2834929   .0627165    -5.70   0.000     .1837523    .4373726
             |
      fixvar |
       Var.  |   .7768361    .082712    -2.37   0.018     .6305208    .9571044
       _cons |   .0851964   .0209345   -10.02   0.000     .0526336    .1379047
------------------------------------------------------------------------------

. predict temp if bkscore==250
(option pr assumed; Pr(sale))
(48 missing values generated)

. replace prob_pred=temp if bkscore==250
(24 real changes made)

. drop temp


. gen pred_profit=ltv*prob_pred
```

Bkscore별 예상되는 LTV 이익률을 내림차순으로 정리하면 아래와 같다.

[bkscore 150]

bkscore	offer	apr	fixvar	fee	people	sale	ltv	prob_pred	pred_profi
150	3	14.9	Fixed	0	4804	0	52	0.0392	2.0384
150	3	14.9	Fixed	0	196	1	52	0.0392	2.0384
150	4	14.9	Var.	0		0	62	0.029453	1.8261
150	4	14.9	Var.	0		1	62	0.029453	1.8261
150	11	19.8	Fixed	0	4928	0	100	0.0144	1.44
150	11	19.8	Fixed	0	72	1	100	0.0144	1.44
150	1	14.9	Fixed	20	4921	0	83	0.0158	1.3114
150	1	14.9	Fixed	20	79	1	83	0.0158	1.3114
150	12	19.8	Var.	0		0	110	0.010751	1.182562
150	12	19.8	Var.	0		1	110	0.010751	1.182562
150	2	14.9	Var.	20	4941	0	93	0.0118	1.0974
150	2	14.9	Var.	20	59	1	93	0.0118	1.0974
150	7	16.8	Fixed	0	4926	0	72	0.0148	1.0656
150	7	16.8	Fixed	0	74	1	72	0.0148	1.0656
150	8	16.8	Var.	0		0	82	0.01105	0.906126
150	8	16.8	Var.	0		1	82	0.01105	0.906126
150	9	19.8	Fixed	20		0	131	0.005716	0.748797
150	9	19.8	Fixed	20		1	131	0.005716	0.748797
150	5	16.8	Fixed	20		0	103	0.005876	0.605251
150	5	16.8	Fixed	20		1	103	0.005876	0.605251
150	10	19.8	Var.	20		0	141	0.004258	0.60036
150	10	19.8	Var.	20		1	141	0.004258	0.60036
150	6	16.8	Var.	20		0	113	0.004377	0.494646
150	6	16.8	Var.	20		1	113	0.004377	0.494646

[bkscore 200]

bkscore	offer	apr	fixvar	fee	people	sale	ltv	prob_pred	pred_profi
200	4	14.9	Var.	0	4791	0	42	0.0418	1.7556
200	4	14.9	Var.	0	209	1	42	0.0418	1.7556
200	2	14.9	Var.	20		0	73	0.021456	1.566295
200	2	14.9	Var.	20		1	73	0.021456	1.566295
200	1	14.9	Fixed	20	4878	0	63	0.0244	1.5372
200	1	14.9	Fixed	20	122	1	63	0.0244	1.5372
200	3	14.9	Fixed	0	4763	0	32	0.0474	1.5168
200	3	14.9	Fixed	0	237	1	32	0.0474	1.5168
200	8	16.8	Var.	0		0	62	0.018635	1.155358
200	8	16.8	Var.	0		1	62	0.018635	1.155358
200	7	16.8	Fixed	0	4894	0	52	0.0212	1.1024
200	7	16.8	Fixed	0	106	1	52	0.0212	1.1024
200	11	19.8	Fixed	0	4936	0	80	0.0128	1.024
200	11	19.8	Fixed	0	64	1	80	0.0128	1.024
200	12	19.8	Var.	0		0	90	0.01124	1.011558
200	12	19.8	Var.	0		1	90	0.01124	1.011558
200	5	16.8	Fixed	20		0	83	0.010769	0.893857
200	5	16.8	Fixed	20		1	83	0.010769	0.893857
200	6	16.8	Var.	20		0	93	0.009454	0.879229
200	6	16.8	Var.	20		1	93	0.009454	0.879229
200	9	19.8	Fixed	20		0	111	0.006475	0.718715
200	9	19.8	Fixed	20		1	111	0.006475	0.718715
200	10	19.8	Var.	20		0	121	0.005681	0.687414
200	10	19.8	Var.	20		1	121	0.005681	0.687414

[bkscore 250]

bkscore	offer	apr	fixvar	fee	people	sale	ltv	prob_pred	pred_profi
250	11	19.8	Fixed	0	4917	0	50	0.017417	0.870851
250	11	19.8	Fixed	0	83	1	50	0.017417	0.870851
250	12	19.8	Var.	0	4928	0	60	0.013583	0.814978
250	12	19.8	Var.	0	72	1	60	0.013583	0.814978
250	2	14.9	Var.	20	4912	0	43	0.018417	0.791932
250	2	14.9	Var.	20	88	1	43	0.018417	0.791932
250	1	14.9	Fixed	20	4878	0	33	0.023583	0.778238
250	1	14.9	Fixed	20	122	1	33	0.023583	0.778238
250	4	14.9	Var.	0		0	12	0.062075	0.744903
250	4	14.9	Var.	0		1	12	0.062075	0.744903
250	8	16.8	Var.	0	4886	0	32	0.0228	0.7296
250	8	16.8	Var.	0	114	1	32	0.0228	0.7296
250	7	16.8	Fixed	0		0	22	0.029159	0.641494
250	7	16.8	Fixed	0		1	22	0.029159	0.641494
250	5	16.8	Fixed	20		0	53	0.008443	0.447464
250	5	16.8	Fixed	20		1	53	0.008443	0.447464
250	6	16.8	Var.	20		0	63	0.006571	0.413972
250	6	16.8	Var.	20		1	63	0.006571	0.413972
250	9	19.8	Fixed	20	4975	0	81	0.005	0.405
250	9	19.8	Fixed	20	25	1	81	0.005	0.405
250	10	19.8	Var.	20		0	91	0.003889	0.353855
250	10	19.8	Var.	20		1	91	0.003889	0.353855
250	3	14.9	Fixed	0		0	2	0.078508	0.157016
250	3	14.9	Fixed	0		1	2	0.078508	0.157016

Round 2

[150]에는 offer 3에, [200]에는 offer 4, [250]에는 offer 11, 12에 나눠서 최종 메일을 보냈다.

Analyze | Final Results

Economics | Round 1 Responses | Round 2 Responses

Round 2 Responses

	150		200		250	
	Sent	Responses	Sent	Responses	Sent	Responses
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	225,000	8,300	0	0	0	0
4	0	0	225,000	10,546	0	0
5	0	0	0	0	0	0
6	0	0	0	0	0	0
7	0	0	0	0	0	0
8	0	0	0	0	0	0
9	0	0	0	0	0	0
10	0	0	0	0	0	0
11	0	0	0	0	117,500	2,164
12	0	0	0	0	117,500	1,924

Cumulative Result

Analyze | Final Results

Round 1 Results

Solicitation Development	$17,000	Round 1 Number Sent	65,000
Mailing Costs	$800	Round 1 Number of Responses	1,487
Cost of Pieces Mailed	$32,500	Total Response Value	$81,562
Total Cost of Mailing	$50,300		
		Total Profit	$31,262

Round 2 Results

		Round 2 Number Sent	685,000
Mailing Costs	$800	Round 2 Number of Responses	22,934
Cost of Pieces Mailed	$342,500	Total Response Value	$1,098,172
Total Cost of Mailing	$343,300		
		Total Profit	$754,872

Cumulative Results

Solicitation Development	$17,000	Round 1 Number of Responses	1,487
Mailing Costs	$1,600	Round 2 Number of Responses	22,934
Cost of Pieces Mailed	$375,000	Total Response Value	$1,179,734
Total Cost of Mailing	$393,600		
		Total Profit	$786,134

[Appendix – Stata code]

```
use exhibit_all, clear

* (1) checking all possible interactions for BK 190 and BK 210 groups from historical data

logistic sale i.apr##i.fixvar i.apr##i.fee i.fee##i.fixvar i.bkscore if bkscore!=255 [fw=people]

* (2) checking all possible interactions for BK 255 group from historical data

logistic sale i.apr##i.fixvar i.apr##i.fee i.fee##i.fixvar if bkscore==255 [fw=people]

* From (1) above, we found that only the interaction between APR and annual fee is meaningful

logistic sale i.apr##i.fee i.fixvar i.bkscore if bkscore!=255 [fw=people]
```

```
* make predictions for only BK 190 and BK 210 groups, not BK 255 group

predict prob_pred if bkscore!=255

* From (2) above, we found that any interactions are not reasonable -> do not use any interactions

logistic sale i.apr i.fee i.fixvar if bkscore==255 [fw=people]

* make predictions for BK 255 group

predict temp if bkscore==255

replace prob_pred=temp if bkscore==255

drop temp


* Using test_result file:

replace people = 4921 in 1

replace people = 79 in 2

replace people = 4941 in 3

replace people = 59 in 4

replace people = 4804 in 5

replace people = 196 in 6

replace people = 4926 in 13

replace people = 74 in 14

replace people = 4928 in 21

replace people = 72 in 22

replace people = 4878 in 25

replace people = 122 in 26

replace people = 4763 in 29

replace people = 237 in 30

replace people = 4791 in 31

replace people = 209 in 32
```

```
replace people = 4894 in 37

replace people = 106 in 38

replace people = 4936 in 45

replace people = 64 in 46

replace people = 4878 in 49

replace people = 122 in 50

replace people = 4912 in 51

replace people = 88 in 52

replace people = 4886 in 63

replace people = 114 in 64

replace people = 4975 in 65

replace people = 25 in 66

replace people = 4917 in 69

replace people = 83 in 70

replace people = 4928 in 71

replace people = 72 in 72

logistic sale i.apr i.fee i.fixvar if bkscore==150 [fw=people]

predict prob_pred if bkscore==150


logistic sale i.apr i.fee i.fixvar if bkscore==200 [fw=people]

predict temp if bkscore==200

replace prob_pred=temp if bkscore==200

drop temp

logistic sale i.apr i.fee i.fixvar if bkscore==250 [fw=people]

predict temp if bkscore==250

replace prob_pred=temp if bkscore==250
```

```
drop temp

gen pred_profit=ltv*prob_pred
```