

Business Context:

In the insurance company, they want to check the renewal rate with a progressive insurance strategy. The progressive insurance strategy consists of 26 factors, including sending letters when cancellation is pending, sending newsletters, and other measures (Figure 1.1). Also, they created 56 combinations of factors, which were written as recipes later (Figure 1.2). They sent the recipes to more than 60,000 policyholders and analyzed the results. Additionally, they divided the customers into four categories based on price change and risk level.

	Factor Description	Predicted Lift
A	Send letter when cancellation is pending.	0.7%
B	Send cancellation notices in special pink envelopes.	0.5%
C	Phone when cancellation is pending.	0.2%
D	Send "stick" letter with renewal quote (no coverage, higher premiums in future).	0.0%
I	Send Progressive merchandise catalog.	0.2%
J	Offer a gift coupon with renewal (mug, tool kit, etc.).	0.9%
K	On renewal quote, put sticker "Policyholder Since 19XX."	0.6%
L	Send "carrot" letter with renewal quote (claim story, accident surcharge waiver, etc.).	-0.1%
M	Include quote stuffer: Here is how you can lower your premium.	0.7%
N	Provide postage-paid envelope with renewal quote.	0.9%
O	Send newsletter.	1.0%
P	Include Progressive magnet with renewal quote.	0.0%
Q	Mail claims comment card to all.	0.3%
R	Provide \$50 deductible coupon with renewal offer.	0.9%
S	Send birthday card (for all names on the policy).	1.4%
T	Include quote stuffer: payment options.	0.9%
U	Send letter after actual lapse.	1.6%
V	Phone after actual lapse.	0.9%
W	Send reminder notice after quote but before expiration.	1.4%
X	Phone with a reminder after quote but before expiration.	0.4%
Y	Send renewal notice in "special" light green envelope.	1.2%

Figure 1.1 Factor Description

	FACTOR																									
Recipe	A	B	C	D	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA
1	1	0	1	1	1	0	0	0	0	0	0	1	0	0	0	1	0	0	1	1	0	1	0	1	0	1
2	1	1	0	1	1	0	0	0	0	0	1	1	0	0	1	0	0	0	1	1	1	1	0	1	1	0
3	0	1	1	1	1	0	0	0	1	0	0	0	1	0	0	1	0	1	0	1	0	1	1	0	1	1
4	0	0	0	1	1	1	1	0	0	1	0	1	0	0	0	1	1	0	1	1	1	1	0	1	0	1
5	0	0	0	1	0	1	1	1	1	0	0	0	0	1	1	0	0	1	1	0	0	1	1	1	1	0
6	0	0	0	0	1	1	1	1	0	1	0	1	0	0	0	1	0	0	1	1	1	1	0	1	1	1
7	1	1	1	0	0	1	0	1	0	0	1	0	0	1	0	1	0	1	0	1	1	1	0	1	1	0
8	1	1	1	0	0	1	1	0	1	0	0	1	0	0	0	0	1	1	1	0	1	1	0	0	1	1
9	1	1	1	0	0	0	1	1	0	1	0	0	1	0	1	0	0	0	1	1	0	1	1	1	0	1
10	0	1	0	0	1	0	0	1	1	1	0	1	0	1	1	0	1	1	0	1	1	1	1	1	0	0
11	0	0	1	1	0	1	0	0	0	1	1	1	1	0	1	1	0	1	1	0	1	1	1	0	0	0
12	1	0	0	0	0	0	1	0	1	0	1	0	1	1	0	1	1	0	1	1	1	1	1	1	0	0
13	0	0	1	0	0	0	0	0	1	1	0	1	1	1	0	1	0	1	0	0	0	1	0	1	1	1
14	1	0	0	0	1	1	0	0	1	1	0	0	1	1	1	1	0	0	0	0	0	1	1	0	1	1
15	0	1	0	1	0	0	1	0	0	1	1	0	1	0	1	0	1	1	0	0	0	0	1	1	1	1
16	0	0	1	0	1	0	1	0	1	0	1	1	0	1	1	1	0	1	1	1	0	0	0	1	0	1
17	1	0	0	1	0	0	0	1	1	1	0	1	1	0	0	1	1	1	1	1	1	0	0	0	1	0
18	0	1	0	0	0	1	0	0	0	1	1	0	1	1	1	0	1	1	1	1	1	0	0	0	0	1
19	1	1	0	1	1	1	0	1	1	0	1	1	1	1	0	0	0	0	1	0	0	0	1	0	0	1
20	0	1	1	1	0	1	1	0	1	1	0	1	1	1	0	0	0	0	0	1	1	0	0	1	0	0
21	1	0	1	0	1	0	1	1	0	1	1	1	1	1	0	0	0	1	0	0	0	1	0	0	1	0
22	1	0	1	1	0	1	0	1	0	0	0	1	0	1	1	1	1	0	0	1	0	1	0	0	0	1
23	1	1	0	0	1	1	1	0	0	0	0	1	1	0	1	1	1	1	0	0	0	0	1	1	0	0
24	0	1	1	1	1	0	1	1	0	0	0	0	1	1	1	1	1	0	1	0	1	0	0	0	1	0
25	1	0	1	1	1	1	1	0	1	1	1	1	1	0	0	0	1	0	1	0	0	1	0	0	1	0
26	1	1	0	1	0	0	1	1	1	1	1	0	0	0	1	1	0	1	0	0	1	0	0	0	0	1
27	0	1	1	0	1	1	0	1	1	1	1	0	0	0	0	1	1	0	1	0	0	1	0	1	0	0

Figure 1.2 Recipe Information with Factors

PROGRESSIVE INSURANCE: MULTIVARIABLE TESTING

The Results

Recipe	PRICE DECREASE				PRICE INCREASE				Total		
	NM		SPU		NM		SPU		Eligible	Renewed	Rate
	Eligible	Renewed	Eligible	Renewed	Eligible	Renewed	Eligible	Renewed			
1	202	148	244	193	295	179	268	190	1009	710	70.4%
2	196	145	232	184	300	202	285	207	1013	738	72.9%
3	206	148	231	190	291	179	284	219	1012	736	72.7%
4	189	130	239	199	304	190	290	211	1022	730	71.4%
5	209	161	216	174	282	174	297	216	1004	725	72.2%
6	200	144	224	185	284	166	298	216	1006	711	70.7%
7	219	154	231	187	273	183	287	209	1010	733	72.6%
8	192	135	230	184	294	188	290	229	1006	736	73.2%
9	212	149	242	184	288	204	295	235	1037	772	74.4%
10	194	139	237	191	296	191	284	204	1011	725	71.7%
11	192	126	234	187	295	193	273	200	994	706	71.0%
12	207	145	240	196	295	185	285	197	1027	723	70.4%
13	205	148	233	194	278	170	264	195	980	707	72.1%
14	193	138	221	179	277	170	302	218	993	705	71.0%
15	176	128	222	178	302	180	292	204	992	690	69.6%
16	191	135	238	194	303	192	289	211	1021	732	71.7%
17	216	154	230	189	285	181	289	212	1020	736	72.2%
18	204	142	238	195	278	177	277	207	997	721	72.3%
19	219	146	237	193	280	167	290	201	1026	707	68.9%
20	222	170	232	183	271	176	287	207	1012	736	72.7%

Figure 1.3 Progressive Insurance Strategy Result

Business Problems:

- 1) Which factors are most likely to contribute to first term renewal?
- 2) Which is the optimal combination of factors based on segmentation level?

Clean the dataset:

We combined the recipe information with progressive insurance results with four segmentations, in order to test the renewal rate.



progressive_revised.csv

Build the model:

In order to test the factors are most likely to contribute the renewal, we are going to build the model with total renewal rate with all factors. Furthermore, we delete all unnecessary variables and only keep the significant variables to run the model again.

```

Call:
lm(formula = Rate ~ D + L + N + U, data = dat)

Residuals:
    Min       1Q   Median       3Q      Max
-0.0314278 -0.0071978  0.0003398  0.0073473  0.0208700

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.712304   0.003654  194.930 < 2e-16 ***
D            -0.007265   0.003268  -2.223 0.030678 *
L            -0.008021   0.003268  -2.454 0.017572 *
N             0.006359   0.003268   1.946 0.057230 .
U             0.011818   0.003268   3.616 0.000685 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01223 on 51 degrees of freedom
Multiple R-squared:  0.353,    Adjusted R-squared:  0.3023
F-statistic: 6.956 on 4 and 51 DF,  p-value: 0.0001507

```

Figure 2.1 Factors Result.

In order to get the optimal combination of factors based on segmentation levels, we will use each segmentation renew rate with all factors.

```

Call:
lm(formula = Rate_1 ~ L, data = dat)

Residuals:
    Min       1Q   Median       3Q      Max
-0.062040 -0.015224 -0.001228  0.017798  0.052739

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.717595   0.005300  135.390 <2e-16 ***
L           -0.014583   0.007496  -1.946  0.0569 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02805 on 54 degrees of freedom
Multiple R-squared:  0.0655,    Adjusted R-squared:  0.0482
F-statistic: 3.785 on 1 and 54 DF,  p-value: 0.05693

```

Figure 3.1 Segmentation1 Result.

```

Call:
lm(formula = Rate_2 ~ K + M + N + Q, data = dat)

Residuals:
    Min       1Q   Median       3Q      Max
-0.041923 -0.006785  0.000000  0.009922  0.030382

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.801044   0.004986 160.663 < 2e-16 ***
K            -0.008455   0.004460  -1.896  0.06365 .
M            -0.008155   0.004460  -1.829  0.07331 .
N             0.009665   0.004460   2.167  0.03492 *
Q             0.012636   0.004460   2.833  0.00658 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01669 on 51 degrees of freedom
Multiple R-squared:  0.2783,    Adjusted R-squared:  0.2217
F-statistic: 4.916 on 4 and 51 DF,  p-value: 0.001981

```

Figure 3.2 Segmentation2 Result

```

Call:
lm(formula = Rate_3 ~ D + U, data = dat)

Residuals:
    Min       1Q   Median       3Q      Max
-0.059835 -0.012979 -0.002358  0.013638  0.063992

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.617940   0.005776 106.981 < 2e-16 ***
D            -0.014812   0.006670  -2.221  0.030666 *
U             0.026402   0.006670   3.958  0.000226 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02496 on 53 degrees of freedom
Multiple R-squared:  0.2799,    Adjusted R-squared:  0.2527
F-statistic: 10.3 on 2 and 53 DF,  p-value: 0.0001663

```

Figure 3.3 Segmentation3 Result

```

Call:
lm(formula = Rate_4 ~ L + Y, data = dat)

Residuals:
    Min       1Q   Median       3Q      Max
-0.051855 -0.013425  0.000451  0.010117  0.061481

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.747239   0.005330 140.204 <2e-16 ***
L            -0.013750   0.006154  -2.234  0.0297 *
Y            -0.012110   0.006154  -1.968  0.0543 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02303 on 53 degrees of freedom
Multiple R-squared:  0.1433,    Adjusted R-squared:  0.111
F-statistic: 4.432 on 2 and 53 DF,  p-value: 0.0166

```

Figure 3.4 Segmentation4 Result

Summary

In this summary table, all significant variables are shown with coefficient. Only the positive coefficient factors can contribute to the renewal rate, therefore, N, Q, and U should be considered in this combination.

Factor	Factor Description	Price decrease		Price increase		Aggregate
		NM	SPU	NM	SPU	
D	Phone when cancellation is pending			-0.01		-0.00727
K	On renewal quote, put sticker "Policyholder Since 19XX"		-0.01			
L	Send "carrot" letter with renewal quote (claim story)	-0.015			-0.01	-0.00802
M	Include quote stuffer: Here is how you can lower premium		-0.01			
N	Provide postage-paid envelope with renewal quote		0.01			0.006359
Q	Mail claims comment card to all		0.013			
U	Send letter after actual lapse			0.026		0.011818
Y	Send renewal letter in "special" light green envelope				-0.01	
	<i>Gain(Implement N,Q,U)</i>	0.00%	2.23%	2.64%	0.00%	1.82%

Figure 3.5 Summary of Different Segmentation Impacts.

Business Insights

The optimal combination, which including postage-paid envelope with renewal quote, mail claims comment card to all and send letter after actual lapse, will contribute to a 1.82% increase in the overall renewal rate. This suggests that the company's strategic efforts are having a positive impact on customer retention and satisfaction.

Furthermore, the optimal combination can have an even greater impact on certain customer segments. For customers whose price is decreasing but who are considered high risk, the optimal combination can increase renewal rates by 2.23%. This indicates that Progressive Insurance's efforts to offer targeted and personalized products and services to high-risk customers are paying off.

For customers whose price is increasing but who are considered low risk, the optimal combination can increase renewal rates by 2.64%. This finding suggests that the company's efforts to communicate the value and benefits of its products and services to low-risk customers are effective.

Overall, these findings provide valuable insights into the effectiveness of Progressive Insurance's pricing and marketing strategies. By identifying specific customer segments and tailoring its products and services to meet their needs, the company is able to improve customer retention and satisfaction, which can ultimately lead to increased profitability and business success.