

# **More on RFM, Assessing a Model's Performance: Gains and Lifts**

**Professor Song Yao**  
Olin Business School

**Customer Analytics**

1

## **How do we conduct RFM in practice?**

---

### **APPLYING RFM (1): TESTING A MARKET CAMPAIGN**

- **Consider discounting test response rates**
  - Reflects fact that most roll outs don't do as well as the tests
  - If discounting by 15%, then a test response rate of 5.00% would be discounted to 4.25%

2

## How do we decide whom to target?

### APPLYING RFM (2): SELECTION WHO TO TARGET

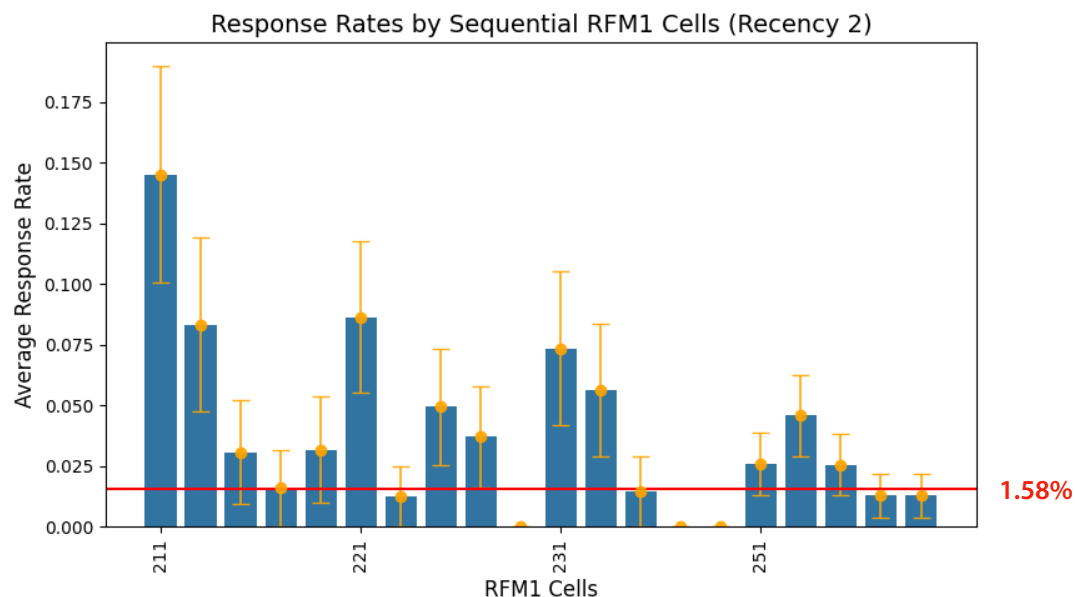
- So far, we have just used the break-even response rate as a strict cutoff
  - **Problem:**
    - The average response rate for a cell is an expected value
    - How do we account for more or less precise predictions?
  - **Solution:**
    - Construct 95% confidence interval of  $p$ , the estimated response rate of a RFM cell
      - ▶ The range within which the truth lies with a 95% probability
- $$p \pm 2 \times \sqrt{\frac{p(1-p)}{n}}, \text{ where } \sqrt{\frac{p(1-p)}{n}} \text{ is the SE of } p$$
- $n$  is the size of the cell (based on the test sample) — another reason why sequential RFM is more preferable.
  - Send the offer if break-even falls **below** the confidence interval (*what is the prob that the true response rate falls below the lower bound?*)

N	p	Plus/Minus	Confidence Interval	
			Lower Bound	Upper Bound
110	0.22	0.078	0.142	0.298
100	0.08	0.054	0.026	0.134
100	0.10	0.060	0.040	0.160

3

## The lower bound of the confidence interval becomes a new measure against which to apply the break-even response rate

### APPLYING RFM (2): SELECTION WHO TO TARGET



4

## Beyond Targeting: What marketing program would you propose for different segments?

---

### APPLYING RFM (3): EXAMPLE

<u>R</u>	<u>F</u>	<u>M</u>	
1	1	4	(1=best, 5=worst)
4	1	1	
5	5	5	
1	5	5	

5

## RFM is very useful in B2B but needs to be modified

---

### RFM IN B2B

- Many fewer customers (often less than 20,000)
- RFM does not beat sales visits -- use for 80% of smaller customers
- 125 cells too many, pick 20-24
- 5 (R), 2 (F), 2 (M) or 2 (R), 2 (F), 5 (M) if monetary is important

6

## RFM works in many other variations

---

### RFM VARIANTS

- Number of N-tiles may vary
  - With relatively small customer databases - quintiles may be too many
  - For very large databases - 125 cells may be too few
  - Don't have to be equal – could have 5 (R), 3 (F), 4 (M) categories
- It is not always RFM
  - E.g., Furniture stores pay little attention to frequency since furniture is a long-lasting product
  - For websites – could be R, F, and D (for duration of visit)

7

### LESSONS ABOUT RFM

1. RFM analysis is
  - Effective
  - Simple
  - Intuitive
  - Flexible
2. Does not require sophisticated software or analytics team  
--> Anyone can do it
3. RFM can handle only few predictive variables
  - There may be much more predictive information

8

## We can assess a model's performance by comparing its performance to the results we expect if no model is used

### TYPES OF MODEL PERFORMANCE MEASURES

- Models are created to predict or classify
- Use model to rank/score customers
- Calculate improvement in response over no targeting

#### Lifts

- "The response rate increase relative to no model"

#### Gains

- "Percentage of total buyers we expect from targeting X% of customers"

9

## To calculate gains we begin with the raw numbers from the Tuango RFM case

### RFM DECILE SUMMARY

```
# Create rfm1_score_decile variable using xtile function
Tuango['rfm1_score_decile'] = xtile(Tuango['avg_response_rate_rfm1'], 10)
Tuango['rfm1_score_decile'] = Tuango['rfm1_score_decile'].max() - Tuango['rfm1_score_decile'] + 1

# Calculate the number of customers and buyers for each decile
decile_summary = Tuango.groupby('rfm1_score_decile').agg(
    total_customers=('userid', 'nunique'),
    total_buyers=('buyer', 'sum')
).reset_index()

print("Customer and Buyer Counts by Score Decile:")
print(tabulate(decile_summary, headers='keys', tablefmt='fancy_grid', showindex=False))

# This is to convert the table to CSV format for easy copy/paste into Excel
decile_summary.to_csv(index=False)
```

Customer and Buyer Counts by Score Decile:

rfm1_score_decile	total_customers	total_buyers
1	1386	144
2	1066	72
3	1662	76
4	1174	40
5	1614	44
6	1427	27
7	1237	14
8	1406	11
9	1171	4
10	1796	0

Export to CSV  
for easy copy/  
paste into excel

Reverse scale so  
that decile 1 has  
best customers

In this example: average response  
probability by rfm cell is the so-  
called "score," namely, the score of  
a customer's buying the product

10

## These raw numbers can be used for the Lift calculations in Excel

### LIFT CALCULATIONS

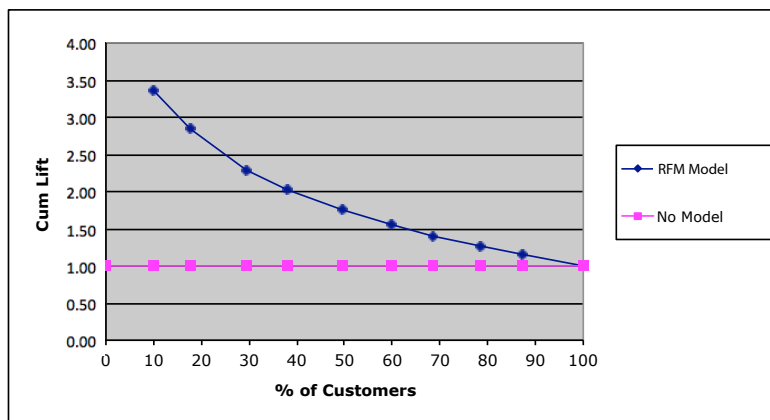
Score Decile	# of Customers	Cum. # of Customers	# of Buyers	Cum. # of Buyers	Resp. Rate by Deciles	Cum. Resp. Rate	Cum. Lift
1	1386	1386	144	144	10.4%	10.4%	3.35
2	1066	2452	72	216	6.8%	8.8%	2.84
3	1662	4114	76	292	4.6%	7.1%	2.29
4	1174	5288	40	332	3.4%	6.3%	2.03
5	1614	6902	44	376	2.7%	5.4%	1.76
6	1427	8329	27	403	1.9%	4.8%	1.56
7	1237	9566	14	417	1.1%	4.4%	1.41
8	1406	10972	11	428	0.8%	3.9%	1.26
9	1171	12143	4	432	0.3%	3.6%	1.15
10	1796	13939	0	432	0.0%	3.1%	1.00
Total	13939		432		3.1%		

- Cumulative # customers: the number of total customers up to and including that decile
- Cumulative # Buyers: the number of buyers up to and including that decile
- Cumulative Response Rate: cumulative # buyers / cumulative # customers
- Cum(ulative) Lift: (cumulative response rate) / (overall response rate)

11

## The Lift indicates the model's ability to beat the "no model"

### CUMULATIVE LIFT CHARTS



- Lift for top decile=3.35: Targeting only these customers we expect to yield 3.35 times the number of buyer than if we did not target
- Note: Lift is relative index, e.g. 3.35 could refer to 4% or 40% response rate

12

## We make the Gains calculations in Excel

### GAINS CALCULATIONS

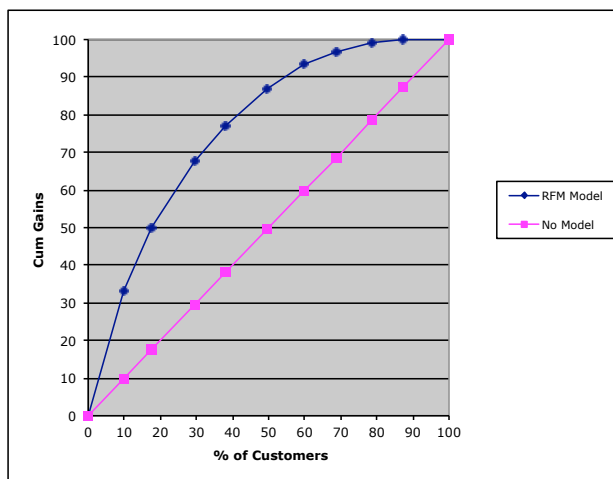
Score Decile	# of Customers	Cum. # of Customers	Cum. % of Customers	# of Buyers	Cum. # of Buyers	Gains	Cum. Gains
1	1386	1386	9.9%	144	144	33.3%	33.3%
2	1066	2452	17.6%	72	216	16.7%	50.0%
3	1662	4114	29.5%	76	292	17.6%	67.6%
4	1174	5288	37.9%	40	332	9.3%	76.9%
5	1614	6902	49.5%	44	376	10.2%	87.0%
6	1427	8329	59.8%	27	403	6.3%	93.3%
7	1237	9566	68.6%	14	417	3.2%	96.5%
8	1406	10972	78.7%	11	428	2.5%	99.1%
9	1171	12143	87.1%	4	432	0.9%	100.0%
10	1796	13939	100.0%	0	432	0.0%	100.0%
Total	13939			432			

- Gains: the proportion of responders (i.e. buyers) in each decile
  - # of Buyers / Total Buyers
- Cum(ulative) Gains: The cumulative gains up to that decile.
  - Cum. # of Buyers / Total Buyers

13

## The Gains chart reveals what proportion of responders we can expect to gain from targeting a specific percent of customers using the model

### CUMULATIVE GAINS CHARTS

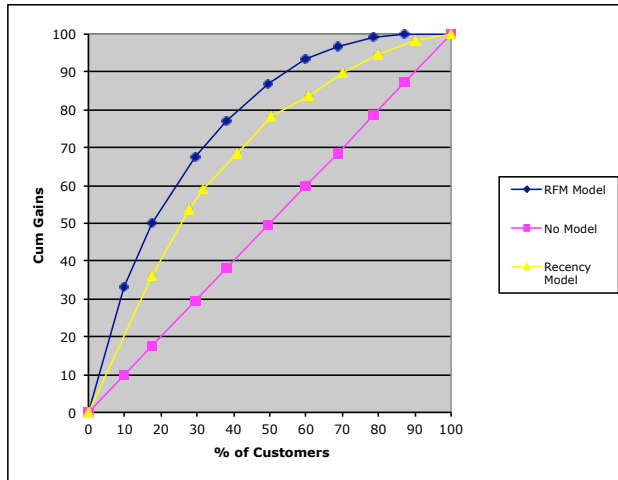


- By using the RFM model to target customers, we can gain 33.3% of buyers by targeting 10% of the customers
- We can gain 76.9% of buyers by targeting 38% of the customers

14

## Lift and Gains can also be used to compare two different models

### EXAMPLE: RECENCY VS. FULL SEQUENTIAL N-TILE RFM MODEL



- The "fatter" the "banana", the better!