

Salad Coupon RFM

User Data Analytics

```
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.1 --
## v ggplot2 3.3.5      v purrr  0.3.4
## v tibble  3.1.4      v dplyr  1.0.7
## v tidyr   1.1.3      v stringr 1.4.0
## v readr   2.0.1      v forcats 0.5.1

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()

library(knitr)
library(kableExtra)

##
## Attaching package: 'kableExtra'
## The following object is masked from 'package:dplyr':
##
##   group_rows

library(flextable)

##
## Attaching package: 'flextable'
## The following objects are masked from 'package:kableExtra':
##
##   as_image, footnote
## The following object is masked from 'package:purrr':
##
##   compose

library(gmodels)
library(Hmisc)

## Loading required package: lattice
## Loading required package: survival
## Loading required package: Formula
##
## Attaching package: 'Hmisc'
## The following objects are masked from 'package:dplyr':
##
##   src, summarize
```

```
## The following objects are masked from 'package:base':
##
##   format.pval, units
library(statar)
library(ggpubr)

##
## Attaching package: 'ggpubr'
## The following objects are masked from 'package:flextable':
##
##   border, font, rotate
salad_coupon <- read.csv("/Users/bowenjin/Desktop/Lion's Choice/rfm_trans.csv")
head(salad_coupon,10)

##      X Card.Number X1...Dollars.Spent.1...Balance X10...Visits.10...Balance
## 1  0  1.741e+11      195.66      15
## 2  1  1.741e+11      562.74      24
## 3  2  1.741e+11       0.00       0
## 4  3  1.741e+11      164.85      13
## 5  4  1.741e+11       0.00       0
## 6  5  1.741e+11      118.63      13
## 7  6  1.741e+11      191.20      25
## 8  7  1.741e+11       0.00       0
## 9  8  1.741e+11      114.63       8
## 10 9  1.741e+11       95.22      11
##      X13...2.Off.Butcher.Block.Sala.13...Redeemed Last.Guest.Activity.Date
## 1      0      31
## 2      0      22
## 3      0      90
## 4      0      24
## 5      0      63
## 6      0      23
## 7      0      42
## 8      0      75
## 9      0      13
## 10     0      29

dim(salad_coupon)

## [1] 25763      6
#response rate of the coupon
CrossTable(salad_coupon$X13...2.Off.Butcher.Block.Sala.13...Redeemed)

##
##
##      Cell Contents
## |-----|
## |              N |
## |      N / Table Total |
## |-----|
##
##
## Total Observations in Table: 25763
```

```
##
##
##      |      0 |      1 |
##      |-----|-----|
##      |  25636 |   127 |
##      |   0.995 |   0.005 |
##      |-----|-----|
##
##
##
##
##
```

Response rate is 0.5%

```
# Create the quintiles for R, F, M
salad_coupon1 <- salad_coupon %>%
  summarise(CardNumber = Card.Number,
            recency = Last.Guest.Activity.Date,
            frequency = X10...Visits.10...Balance,
            monetary = X1...Dollars.Spent.1...Balance,
            coupon = X13...2.Off.Butcher.Block.Sala.13...Redeemed,
            rec_quin = xtile(Last.Guest.Activity.Date, 5),
            freq_quin = xtile(X10...Visits.10...Balance, 5),
            mv_quin = xtile(X1...Dollars.Spent.1...Balance, 5))
head(salad_coupon1,10)
```

```
##      CardNumber recency frequency monetary coupon rec_quin freq_quin mv_quin
## 1  1.741e+11     31      15    195.66      0      2      4      4
## 2  1.741e+11     22      24    562.74      0      2      5      5
## 3  1.741e+11     90       0     0.00      0      4      1      1
## 4  1.741e+11     24      13    164.85      0      2      4      4
## 5  1.741e+11     63       0     0.00      0      3      1      1
## 6  1.741e+11     23      13    118.63      0      2      4      4
## 7  1.741e+11     42      25    191.20      0      3      5      4
## 8  1.741e+11     75       0     0.00      0      4      1      1
## 9  1.741e+11     13       8    114.63      0      2      4      4
## 10 1.741e+11     29      11     95.22      0      2      4      4
```

```
#check and adjust ranking for R, F, M
```

```
#Recency rank
```

```
salad_coupon1 %>% group_by(rec_quin) %>% summarise(avg_rec = mean(recency), .groups="drop")
```

```
## # A tibble: 5 x 2
##   rec_quin avg_rec
##   <int>   <dbl>
## 1     1    5.34
## 2     2   19.9
## 3     3   48.5
## 4     4   91.6
## 5     5  156.
```

```
#Frequency rank adjust
```

```
salad_coupon1 %>% group_by(freq_quin) %>% summarise(avg_freq = mean(frequency), .groups="drop")
```

```
## # A tibble: 4 x 2
##   freq_quin avg_freq
```

```
##           <int>      <dbl>
## 1           1         0
## 2           3      2.84
## 3           4     10.2
## 4           5     48.4

salad_coupon1$freq_quin <- max(salad_coupon1$freq_quin) + 1 - salad_coupon1$freq_quin
salad_coupon1 %>% group_by(freq_quin) %>% summarise(avg_freq = mean(frequency), .groups="drop")

## # A tibble: 4 x 2
##   freq_quin avg_freq
##       <dbl>   <dbl>
## 1         1     48.4
## 2         2     10.2
## 3         3      2.84
## 4         5         0

#Monetary rank adjust
salad_coupon1 %>% group_by(mv_quin) %>% summarise(avg_mv = mean(monetary), .groups="drop")

## # A tibble: 4 x 2
##   mv_quin avg_mv
##       <int> <dbl>
## 1         1         0
## 2         3     29.4
## 3         4    123.
## 4         5   571.

salad_coupon1$mv_quin <- max(salad_coupon1$mv_quin) + 1 - salad_coupon1$mv_quin
salad_coupon1 %>% group_by(mv_quin) %>% summarise(avg_mv = mean(monetary), .groups="drop")

## # A tibble: 4 x 2
##   mv_quin avg_mv
##       <dbl> <dbl>
## 1         1   571.
## 2         2    123.
## 3         3     29.4
## 4         5         0

#create rfm index
salad_coupon1 <- salad_coupon1 %>%
  mutate(rfmindex_iq = 100*rec_quin + 10*freq_quin + mv_quin)

head(salad_coupon1,10)

##   CardNumber recency frequency monetary coupon rec_quin freq_quin mv_quin
## 1  1.741e+11     31       15    195.66      0         2         2      2
## 2  1.741e+11     22       24    562.74      0         2         1      1
## 3  1.741e+11     90        0      0.00      0         4         5      5
## 4  1.741e+11     24       13    164.85      0         2         2      2
## 5  1.741e+11     63        0      0.00      0         3         5      5
## 6  1.741e+11     23       13    118.63      0         2         2      2
## 7  1.741e+11     42       25    191.20      0         3         1      2
## 8  1.741e+11     75        0      0.00      0         4         5      5
## 9  1.741e+11     13        8    114.63      0         2         2      2
## 10 1.741e+11     29       11     95.22      0         2         2      2
##   rfmindex_iq
```

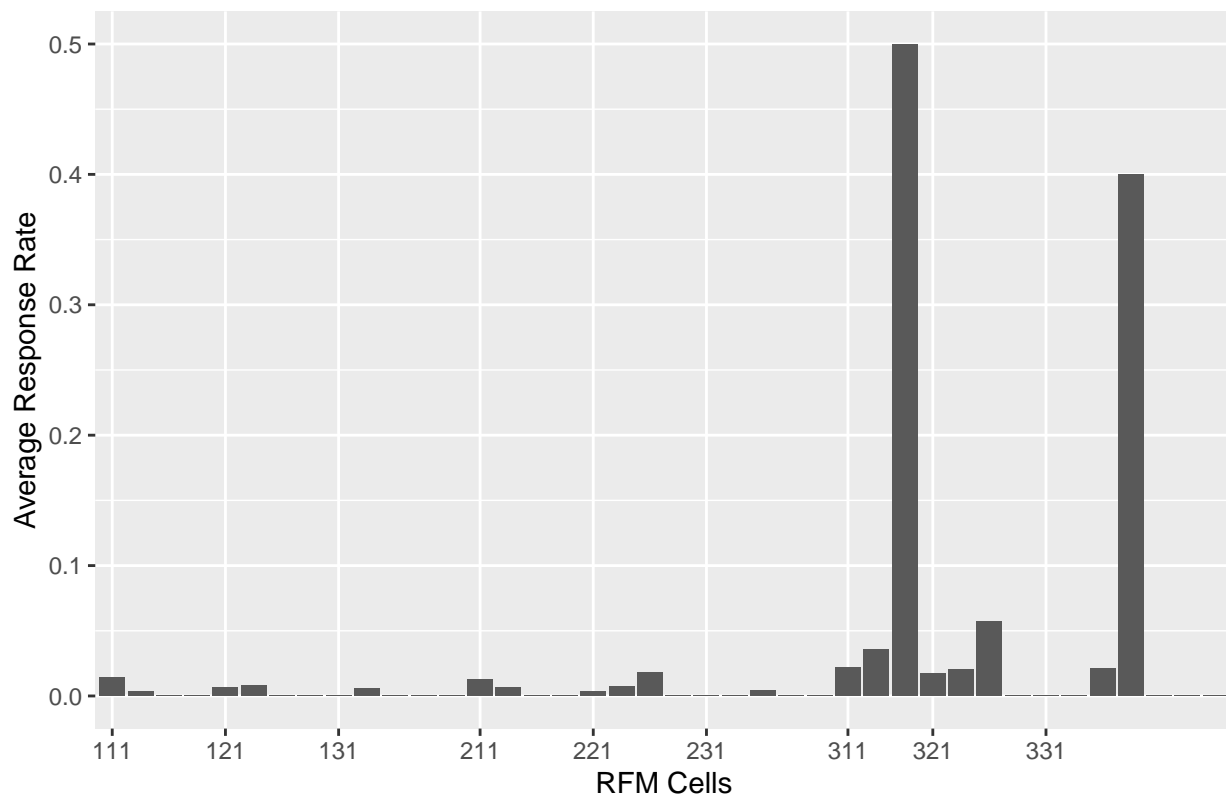
```
## 1      222
## 2      211
## 3      455
## 4      222
## 5      355
## 6      222
## 7      312
## 8      455
## 9      222
## 10     222
```

```
#response rate in each RFM group
avg_resp_rate_rfm <- salad_coupon1 %>%
  group_by(rfmindex_iq) %>%
  summarise(resp_rate_rfm_iq=mean(coupon), .groups="drop") %>%
  arrange(desc(resp_rate_rfm_iq))
head(avg_resp_rate_rfm,10)
```

```
## # A tibble: 10 x 2
##   rfmindex_iq resp_rate_rfm_iq
##       <dbl>         <dbl>
## 1         315             0.5
## 2         335             0.4
## 3         323            0.0571
## 4         312            0.0357
## 5         311            0.0215
## 6         333            0.0208
## 7         322            0.0206
## 8         223            0.0180
## 9         321            0.0174
## 10        111            0.0143
```

```
bar_avg_resp_rate_rfm <-
  ggplot(data=avg_resp_rate_rfm,
    aes(x = as.factor(rfmindex_iq), y = resp_rate_rfm_iq)) +
  labs(x="RFM Cells",
    y="Average Response Rate",
    title = "Response Rates by Independent RFM Cells") +
  theme(plot.title = element_text(hjust = 0.5)) +
  geom_bar(stat="identity") +
  scale_x_discrete(breaks = seq(111, 555, by = 5))
bar_avg_resp_rate_rfm
```

Response Rates by Independent RFM Cells



```
#response rate for every member
salad_coupon1 <- salad_coupon1 %>%
  group_by(rfmindex_iq) %>%
  mutate(resp_rate_by_rfm_iq = mean(coupon)) %>% ungroup()

head(salad_coupon1,10)
```

```
## # A tibble: 10 x 10
##   CardNumber recency frequency monetary coupon rec_quin freq_quin mv_quin
##   <dbl>      <int>      <dbl>      <dbl> <dbl>      <int>      <dbl>      <dbl>
## 1 174100000028     31        15      196.      0         2         2         2
## 2 174100000176     22        24      563.      0         2         1         1
## 3 174100000440     90         0        0        0         4         5         5
## 4 174100000515     24        13      165.      0         2         2         2
## 5 174100000770     63         0        0        0         3         5         5
## 6 174100000903     23        13      119.      0         2         2         2
## 7 174100001000     42        25      191.      0         3         1         2
## 8 174100001166     75         0        0        0         4         5         5
## 9 174100001216     13         8      115.      0         2         2         2
## 10 174100001687     29        11      95.2      0         2         2         2
## # ... with 2 more variables: rfmindex_iq <dbl>, resp_rate_by_rfm_iq <dbl>
```