

Find the Best Consumers for Monro

Fabulous 1



Overview

FIND THE BEST CURRENT CONSUMERS

Based on RFM metrics

- Divide consumers based on a single variable.
- Define each individual consumers into segments by rating.



Build individual table on RFM for further analysis



IndividualID

Recency

Frequency

Monetary

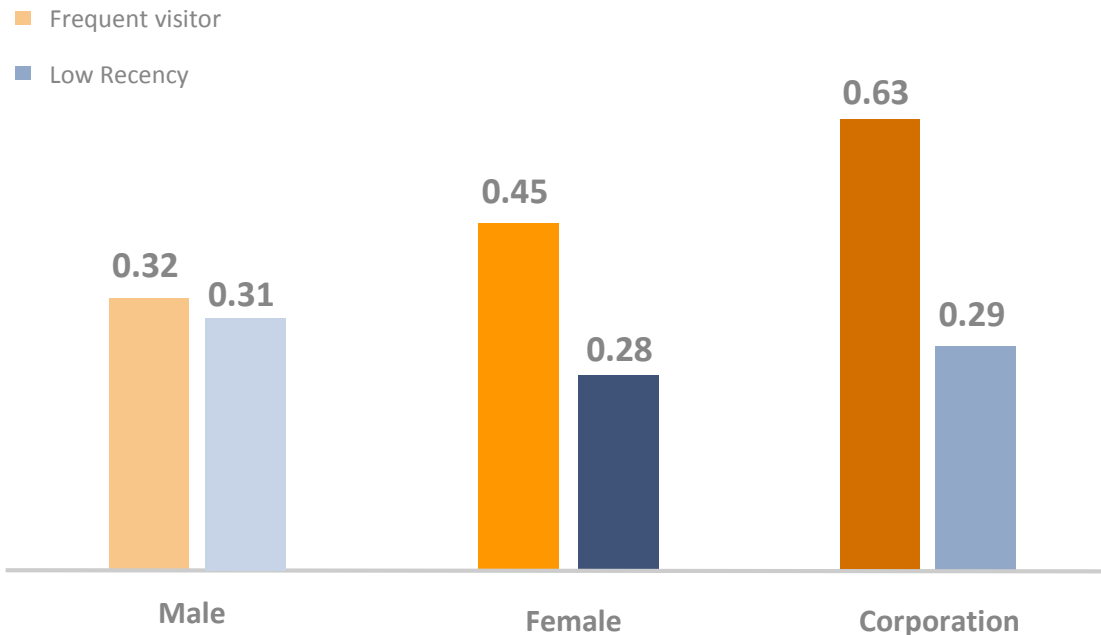
Duration

Average
Spending

- **Recency** : Time since last order (day)
- **Frequency** : Total number of transactions (#)
- **Monetary** : Total sales amount (\$)
- **Duration** : Time from last purchase to first purchase
- **Average Spending**: $\text{Monetary} / \text{Frequency}$ (\$)



Segmentation based on a single variable (Frequency or Recency) is not accurate



- **Frequent visitor:** who visited us more than once a year.
- **Low recency:** who visited us within a year

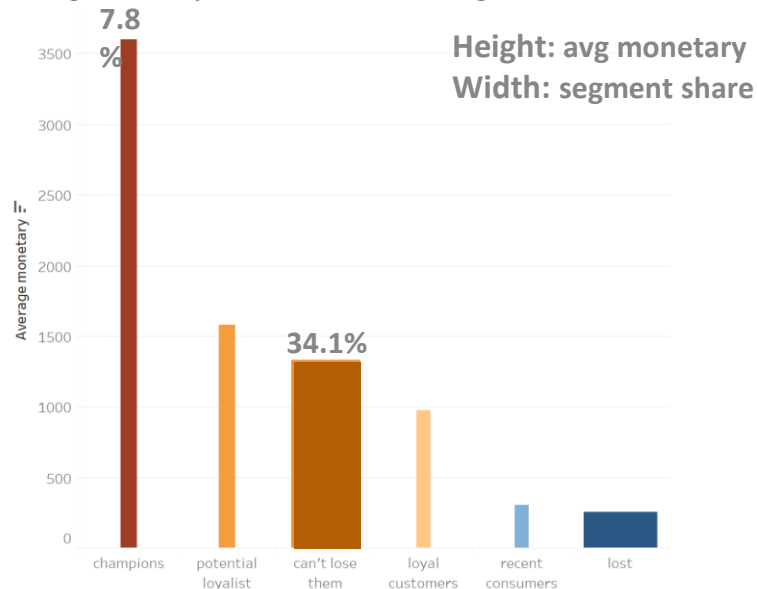


The number of can't lose group (34.1%) outweigh the number of ideal consumers (7.8%)

	DummRec	Avg Spend	FrevalueDM
Champions	✓	✓	✓
PotentLoyal	✓	✓	
LoyalCus	✓		✓
RecentCus	✓		
Can't lose		✓	(✓)
Lost			(✓)

Already lost

Average monetary and customer size in segments



*segment details are in appendix 3

✓: short recency, high average spending, high frequency



Customer Rating based on RFM can value each segments, as well as future target marketing

IndividualID

Recency

Frequency

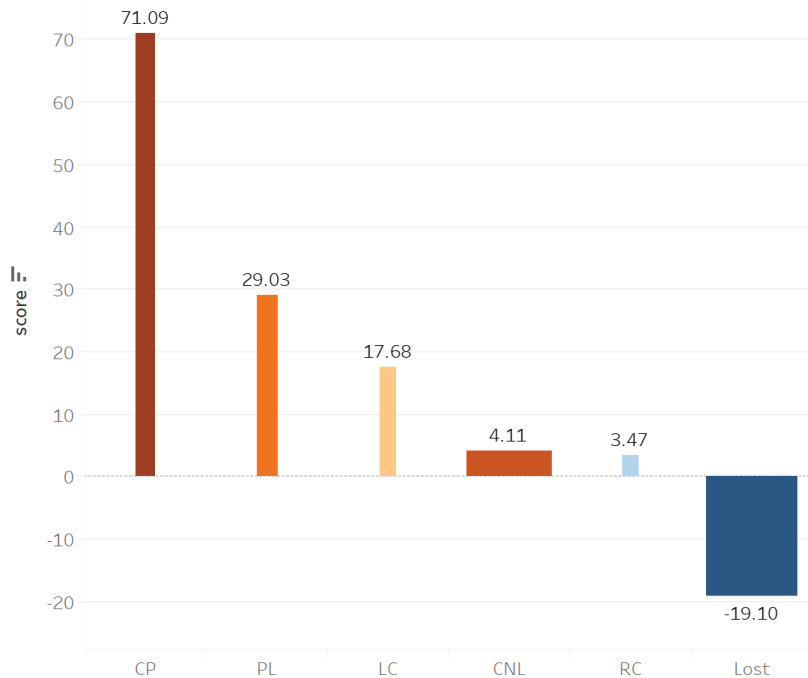
Monetary

Duration

Average
Spending

Score

$$\text{Score} = (3.5 * \text{Frequency} + 0.05 * \text{Monetary} - 0.05 * \text{Recency}) / 25$$



Champions

PotentLoyal

LoyalCus

Can't lose

RecentCus

Lost

Best consumers

Consider winning back

Appendix



Appendix1: How we built up RFM table for each individual from choice data

	IndividualId	first_purchase	last_purchase	duration	frequency	monetary	recency	avg_spending	score
1	1000000000257	2015-01-06	2015-01-06	0	1	26.95	1852	26.95000	-3.51010
2	1000000000483	2016-02-24	2016-07-27	154	2	489.61	1284	244.80500	-1.30878

duration

```
first_purchase=aggregate(InvoiceDate~IndividualId,FUN=min,data=monro)
last_purchase=aggregate(InvoiceDate~IndividualId,FUN=max,data=monro)
individual=merge(first_purchase,last_purchase,by='IndividualId')
names(individual)<- c('IndividualId','first_purchase','last_purchase')
individual$duration=individual$last_purchase-individual$first_purchase
```

frequency

```
frequency=aggregate(InvoiceId~IndividualId,FUN=length,data=monro)
individual=merge(individual,frequency,by='IndividualId')
```

monetary

```
monetary=aggregate(GrossSalesAmount~IndividualId,FUN=sum,data=monro)
individual=merge(individual,monetary,by='IndividualId')
names(individual)<-c('IndividualId','first_purchase','last_purchase','duration','frequency','monetary')
individual$recency= as.numeric(as.Date('2020-02-01')-individual$last_purchase)
individual$avg_spending=individual$monetary/individual$frequency

individual=individual[individual$monetary>0,]
individual$score<-(3.5*individual$frequency+0.05*individual$monetary-0.05*individual$recency)/25
```




Appendix2 : Regression on gender and frequency and recency

```
> glm(as.factor(frevalueDM)~0+Gender,gender,family  
= "binomial")
```

Coefficients:

GenderF	GenderM	GenderU
0.4483	0.3245	0.6268

Degrees of Freedom: 39591 Total (i.e. Null); 39588 Residual

Null Deviance: 54880

Residual Deviance: 53360 AIC: 53360

```
> summary(lm(recency~0+Gender,gender))
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
GenderF	930.801	6.237	149.25	<2e-16 ***
GenderM	884.667	4.492	196.92	<2e-16 ***
GenderU	908.619	12.441	73.03	<2e-16 ***

Residual standard error: 696 on 39588 degrees of freedom

Multiple R-squared: 0.6264, Adjusted R-squared: 0.6264

F-statistic: 2.213e+04 on 3 and 39588 DF, p-value: < 2.2e-16



Appendix3 : How we define dummy variables for RFM and make consumer segments

```
oldcustomer$frevalue<-oldcustomer$frequency / as.numeric(oldcustomer$duration)
summary(oldcustomer$frevalue)
```

```
##      Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
## 0.0007843 0.0038596 0.0068120      Inf 0.0123457      Inf
```

```
oldcustomer$frevalueDM<-oldcustomer$frevalue>2/365 # 2 times in a year
oldcustomer$frevalueDM[oldcustomer$frevalueDM==TRUE] <- 'highfrevalue'
oldcustomer$frevalueDM[oldcustomer$frevalueDM==FALSE] <- 'lowfrevalue'
table(oldcustomer$frevalueDM)
```

```
##
## highfrevalue lowfrevalue
##      24261      16109
```

```
oldcustomer$dummRecen <- oldcustomer$recency < 365
oldcustomer$dummRecen[oldcustomer$dummRecen==TRUE] <- 'shortRec'
oldcustomer$dummRecen[oldcustomer$dummRecen==FALSE] <- 'longRec'
```

```
summary(oldcustomer$avg_spending)
```

```
##      Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
##      0.334   49.780  133.228  187.920  263.202  4173.550
```

```
oldcustomer$avgSpend <- oldcustomer$avg_spending >133
oldcustomer$avgSpend [oldcustomer$avgSpend ==TRUE]<- 'highSpend'
oldcustomer$avgSpend [oldcustomer$avgSpend ==FALSE]<- 'lowSpend'
```

```
dt = data.table(oldcustomer)
dagg = dt[,.(monetary=mean(monetary), number=length(monetary)), by = .(avgSpend,frevalueDM,dummRecen)]
dagg
```

```
##      avgSpend  frevalueDM dummRecen  monetary number
## 1: highSpend highfrevalue   longRec 1521.4314   8770
## 2: lowSpend  highfrevalue   longRec  310.1774   9597
## 3: lowSpend  highfrevalue  shortRec  954.4399   2743
## 4: highSpend lowfrevalue   shortRec 1582.0341   3305
## 5: highSpend highfrevalue  shortRec 3605.3926   3151
## 6: highSpend lowfrevalue   longRec 1142.3683   4986
## 7: lowSpend  lowfrevalue   shortRec  308.9817   2754
## 8: lowSpend  lowfrevalue   longRec  202.4866   5064
```

```
##      group churn_rate  groupavg  purchfre  purchrec  purchmon  Num
## 1      CP  0.2097747  278.72952 14.010790 149.1815 3605.3926 3151
## 2      PL  0.2580938  330.93160  5.051740 165.9973 1582.0341 3305
## 3      CNL 1.0000000  322.36673  4.650989 1211.1038 1384.0362 13756
## 4      LC  0.2511848  68.28011 12.770689 159.6602  954.4399 2743
## 5      RC  0.2614379  60.68209  4.863471 169.6852  308.9817 2754
## 6      Lost 1.0000000  56.30058  4.432985 1258.8242  272.9803 14661
##      CLTV      score
## 1 18616.2645  71.085733
## 2  6477.4139  29.027980
## 3  1499.3240   4.109786
## 4  3471.4836  17.683491
## 5  1128.8555   3.466817
## 6   249.5797 -19.096259
```

Find Best Consumers for Monro

Lujie Chen | Shuyi Chen | Sahyun Kim | Fan Yang | Jinglin Zhu

Fabulous 1