

Retail Relay Revisited

Introduction

Because Retail Relay (Relay) was operating like a start-up—all employees scrambling around to get everything done in the face of rapid growth—some important business processes had been left unattended. One of the processes that received little attention was how to use the customer-level purchase information in Relay’s database to improve customer retention.

Relay understood that many of its existing customers were not only spending money with Relay, but were also purchasing some of their grocery products from other vendors as well. This kind of customer behavior was evident from even a casual examination of the customer-level purchase data. Some customers purchased from Relay infrequently and sporadically, so clearly, these customers must be shopping somewhere else during the interludes between their purchases from Relay. Because customers did not sign up for a subscription plan, Relay could never be certain if customers stopped purchasing from Relay (or “churned”) or if they were merely dormant for a while. To overcome this challenge, Relay used a rule of thumb that classified customers as churned if their dormancy duration was more than two standard deviations above the customer’s mean interpurchase time. For example, consider a customer with a mean interpurchase time of two months and the standard deviation in interpurchase time of three months. This customer is classified as churned if his or her dormancy duration is more than eight months.

Churned customers represented a loss of potential profit, and management believed that this loss was substantial. Yet Relay had not taken the time and energy to fully leverage its customer-level transaction data—an important customer-relationship asset—to improve customer retention. Because customers submitted their orders through Relay’s website, Relay had a large and detailed database of

customer orders. Among other information, when a customer placed an order, Relay knew which customer made the order, what items that customer ordered, the customer's entire order history, and the date and time of these orders. Relay decided to use the customer-level purchase data to better understand the factors that influenced customer retention. Were there any distinct characteristics of the retained customers that could instruct Relay about how to increase the retention of its not-so-regular customers?

Relay management knew that somewhere in this data lay the key to unlocking more of its current customers' grocery dollars. Now, it needed to dig until it found that key. Would a logistic regression analysis reveal the keys to improving customer retention? To begin this process, Relay decided to explore the customer summary file described in Exhibit 14-1.

Exhibits

Exhibit 14-1 Retail Relay: Description of Customer Summary File

Variable	Description
custid	Computer-generated ID to identify customers throughout the database
retained	1 if customer is assumed to be retained; 0 otherwise
created	Date when the contact was created in the database—when the customer joined
firstorder	Date when the customer placed first order
lastorder	Date when the customer placed last order
esent	Number of e-mails sent
eopenrate	Number of e-mails opened divided by number of e-mails sent
eclickrate	Number of e-mails clicked divided by number of e-mails sent
avgorder	Average order size for the customer
ordfreq	Number of orders divided by customer tenure
paperless	1 if customer subscribed for paperless communication (only online)
refill	1 if customer subscribed for automatic refill
doorstep	1 if customer subscribed for doorstep delivery
train	1 if customer is in the training database
favday	Customer's favorite delivery day
city	City where the customer resides

Source: Company documents.