If Only I Knew…

By Tim Graettinger

If only I knew …
- who will renew their service plan
- who will defect to the competition
- who will buy our new product
- who will become a high-value customer

Sound familiar? To make better decisions as business people, we all wish we could see into the future and deep into the hearts and minds of our customers. In response to the wishes above, this article asks the question, "Suppose I knew …, what would I do?" We will describe the process to translate the wish for customer knowledge into actionable strategies and tactics.

We will look specifically at one common business problem, customer retention, but the same principles will apply to a broad range of business issues. The primary instruments involved will be a customer database/warehouse and a predictive model built via data mining. By the time you are finished reading, you will understand the wish-to-action translation process from end to end. And you will be able to apply the process to your own business problem(s) – before any money is spent or any project is begun.

To keep the discussion concrete, we will consider the process in the context of a case study. Suppose that Donna is the VP of Marketing for a large trade organization[1]. She is responsible for several trade shows and a large annual meeting. Over time, there has been a decrease in attendance at the annual meeting. She needs to increase retention. Donna asks, "Suppose I knew who will come to this year's meeting, what would I do?"

We will continue Donna's exercise of imagination by:
- mapping the territory,
- segmenting the territory into groups or clusters, and
- designing strategies and tactics for each group

Mapping the Territory

First, consider Donna's assumption, "Suppose I knew who will come to this year's meeting." Determining who will come is not as simple as querying the customer database. Why? The customer database can only reveal who came last year, but not who will come this year. More generally, a customer database or warehouse can report what happened, but not what will happen next. To look ahead rather than look back, data mining[2] is an excellent option.

Data mining is the discovery and modeling of hidden patterns in large volumes of data. It differs from database querying and reporting in that the data mining process produces a model. A data mining model can take the form of a set of "if-then" rules or a mathematical formula – either of which can be generated heuristically by Donna, by semi-automated statistical means, or by a combination of both[3]. The model uses data about the past and present to predict future outcomes. In Donna's situation, relevant historical data about a prior attendee might include: how many times they attended previously, their age, their industry, and the size of their company. The model outcome is a prediction about whether or not they will come to this year's meeting.

Such models have technical labels like predictive models, likelihood models, or scoring models. The names are actually quite descriptive for the current discussion. For Donna, a typical likelihood model will produce a score, say from 0 to 100, indicating how likely an individual is to attend this year's meeting. The larger the score is, the more likely the outcome is. One individual might score high because they have attended for many years and they are in the "Baby Boomer" age group. Another individual might have attended just once, resulting in a lower score. Conceptually, Donna's prior attendees are shown ranked from low to high in Figure 1. Each dot or mark represents a prior attendee, and we can see that some are highly likely to attend this year, while others are much less so.

Figure 1 – Ranking of Prior Attendees by Likelihood to Attend

At this point, it is tempting to think that this is a sufficient (conceptual) map of the customer landscape. Others have jumped into data mining efforts with even less reflection than this. Why not jump ahead? Simply put, not all attendees or customers are created equal. To realize the full benefit of data mining for increasing revenue, reducing cost, and/or improving ROI, we need some notion of customer value. Then the territory map will be truly actionable.

Different attendees at the annual meeting certainly have different values to Donna's company. Some register for the full, multi-day meeting. Others attend only for a single day. Certain attendees belong to sponsor organizations, those who pay for promotional considerations. Some are exhibitors who pay for booth space to display their products and talk
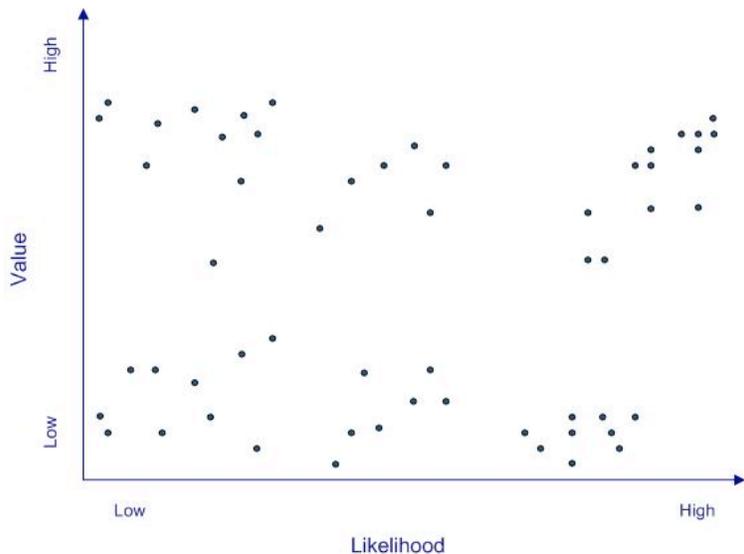

Figure 2 – Map of Prior Attendees by Likelihood to Attend and Value

with existing and potential customers. Others have intrinsic value to the meeting, say as speakers or volunteers. The variations of value go on and on.

In general, a second predictive model is necessary to estimate the likely value of a customer for this year's meeting. For simplicity of discussion, however, we will use a simple estimate of attendee value. The estimate will be the average of their attendance values at previous meetings, ignoring any intrinsic worth for now. Adding this new value dimension to the likelihood score, we produce the final map of the territory, shown in Figure 2. As before, each mark on the map is a prior attendee. Notice that many different combinations of likelihood and value are present. Some prior attendees are high value and highly likely to attend. Others are high value but less likely to attend, and so on. The picture has become slightly more complex, but it needs to be to make good decisions.

Segmenting the Territory

Having assigned a likelihood of attendance and a likely value to each prior attendee, we have addressed the first half of Donna's question. Now, we move forward to consider the second half, "what would I do?"

For marketing and many other business functions, it is common not to work with the precise estimates for likelihood and value for each customer. Rather, it often makes sense to group individuals into segments or clusters with similar attributes of likelihood and value. Donna can define marketing strategies and tactics that fit the common characteristics of the segment, rather than try to customize to each unique individual.

Figure 3 depicts a simple conceptual segmentation of the likelihood-value map. We split the territory into four blocks that represent useful, distinct combinations: low likelihood/low value (LL) in the southwest quadrant, high likelihood/low value (HL) in the southeast quadrant, low likelihood/high value (LH) in the northwest quadrant, and high likelihood/high value (HH) in the northeast quadrant. We will give these segments more descriptive names later in the article.
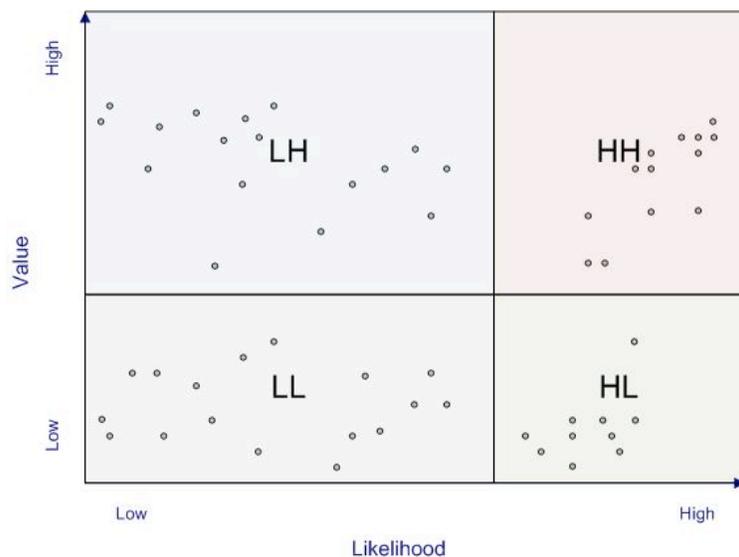

Figure 3 – Segmented Map of Prior Attendees

Segmenting the territory into more or fewer than four segments is not unusual. The choice for the number of segments can be driven by any natural clustering that is evident. Or, the choice can depend on the need to keep the number of segments at a manageable level. For Donna's purposes, the four segments illustrated in Figure 3 will suffice.

Designing Strategies and Tactics

Now, Donna can get creative. With four segments defined, she can think strategically about how she wants to develop the attendees in each segment. Refer to Figure 4 to visualize the discussion that follows. For instance, strategically, Donna wants to keep the "Sweet Spot" (HH) attendees where they are. To achieve that goal, she considers tactics like a loyalty/rewards program or special opportunities to meet with thought leaders at the annual meeting.

The "Frequent but Frugal" (HL) segment consists of individuals who are likely to attend, but of low value. Donna's strategic goal is to move them into the "Sweet Spot" portion of the map. Her tactics can include creating incentives for these prior attendees to attend the full conference or to become sponsors or exhibitors - that is, to become high-value attendees.
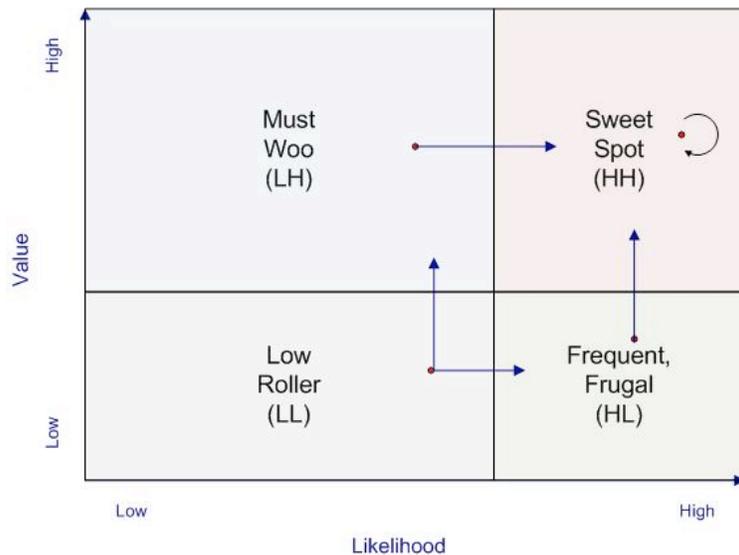

Figure 4 – Strategic Movement Map for Prior Attendee Segments

Conversely, the "Must Woo" (LH) segment includes individuals who are unlikely to attend, but of high value when they do. Again, Donna's strategic goal is to move them into the "Sweet Spot" quadrant. Tactically, she needs to provide incentives for these people to attend – in other words, to increase their likelihood.

Last, Donna considers the "Low Roller" (LL) segment. Since these prior attendees are neither high likelihood nor high value, she decides to use them to test different marketing messages. Randomly, one Low Roller test group will get the Frequent but Frugal message, trying to improve the value of their attendance. A different Low Roller test group will get the Must Woo message, attempting to improve the likelihood of attendance. By virtue of testing, Donna can learn which is more successful: moving Low Roller individuals into the Must Woo

segment or moving them into the Frequent but Frugal segment. From there, she can apply the other tactics described above.

Conclusion

As stated at the outset, the goal of this article was to describe a process to translate the wish for customer knowledge into actionable strategies and tactics. We illustrated the process using a retention-oriented case study, although the process is more generally applicable. Along the way, we mapped the customer territory, combining estimates of both likelihood and value. The likelihood estimate was created via a data mining model, while the value estimate was extracted directly from the customer database. Finally, the territory was segmented into distinct groups, with strategies and tactics designed for each one.

All of this was done in the mind's eye, conceptually, before a single dollar was spent. Perhaps best of all, a sequence of simple pictures (like Figures 1-4) illustrates the process, the expected outcomes from it, and how it addresses the business problem at hand. We all know that a sequence of simple pictures is very valuable in making a business case when it comes time to devote money and resources – resources that will make an "if I only knew…" wish come true.

Tim Graettinger, Ph.D., is the President of Discovery Corps, Inc. (http://www.discoverycorpsinc.com), a Pittsburgh-area company specializing in data mining, visualization, and predictive analytics. Contact Tim at (724)-743-3642 or tgraettinger@discoverycorpsinc.com

---

[1] Donna was introduced in a previous article. See "Data Mining Questions? Some Back-of-the-Envelope Answers" at http://www.discoverycorpsinc.com
[2] For more background on data mining, see "Digging Up Dollars with Data Mining" at http://www.discoverycorpsinc.com
[3] For additional information about the model-building process, again see "Digging Up Dollars with Data Mining" at http://www.discoverycorpsinc.com