

Simplified Data Analytics for the Accurate Evaluation of a New Venture's Market Potential

Michael D. Mattei
Stephen Hellebusch

This article examines the creation of an accurate market projection designed with easy-to-use, cost-effective data analytic techniques. Many of the techniques explored are derived from the subdisciplines of decision support and data warehousing found in the information technology arena. Two significant contributions are presented: a simple mathematical technique that eliminates the need for heuristics, and the simplification of the process to the point where no computer or sophisticated statistical analysis is needed.

One question that continues to challenge entrepreneurs is how to determine the market potential for a new product or service, especially a product or service that is unique in the marketplace. Many view the process as putting some numbers on a sheet of paper and calling it a sales forecast. Investors, bankers, venture capitalists, even "family, friends and fools," often view such projections more as wishful thinking than a reliable estimate.

The focus of this article is on the creation of an accurate market projection or forecast developed with easy-to-use and cost-effective data analytic techniques. Many of the techniques presented are drawn from the subdisciplines of decision support and data warehousing found in the information technology (IT) arena. IT practitioners use terms that seem quite foreign to entrepreneurs and businesspeople, but are easily translated into common business nomenclature as shown below.

- **IT version.** The process presented is built upon the collection of primary data, creating a data mart, merging it with secondary data, applying data analytic tools, and drilling down to provide information for decision support.
- **Translation.** The process presented is built upon the collection of data specific to the venture (e.g., a concept survey), creating a table of venture specific data, merging it with readily available nonventure specific census, industry or trade data, analyzing the data, and looking at various slices of the data to create insights that provide an accurate and reliable assessment of the market potential for a new product or service.

The starting point of the data analytic process described here is the concept survey. These surveys are widely used by large organizations but are expensive to field and diffi-

cult to analyze. Analysis of the data often requires elaborate mathematical or statistical techniques. This traditional approach to assessing market potential can be effective, but often too expensive, too complex, or both. To further complicate the traditional approach, the data often need to be modified using heuristics to adjust for exaggerated purchase intent.

Two significant contributions are made in this article. First, a simple mathematical technique is shown that eliminates the need for heuristics (or rules of thumb) often used to adjust data for overstated purchase intent. The second contribution is the simplification of the process to the point that it is possible to execute without the use of a computer or sophisticated statistical analysis, although it may take a little more time. In the end, whether using a computer or not, the value of the information obtained will far exceed the cost of acquiring and manipulating the data.

Which Forecasting Approach Is Best for a New Venture?

Most, if not all new ventures begin as a great idea. A great idea can become a great opportunity if there is sufficient market demand. There are four general approaches to forecasting demand for a product or service. The first uses historical data to project the future. The past trend is extended to future time periods as an estimate of sales. For new ventures, especially the truly innovative ones, there are usually no historical data so this approach is ineffective.

A second approach uses a "life-cycle" assumption, which appears like an S-curve. The life-cycle theory of the firm has been discussed widely in the management, marketing, and economics literature and generally labels the four stages as introduction, growth, maturity, and decline. The drawback of this approach is that the initial point is extremely difficult to determine, as is the length of time for each of the four stages. If this approach is used, the initial point is usually an educated guess or an estimate of a percentage of a target population. For example, a local entrepreneur who had departed from a very successful senior corporate position estimated that his new venture would obtain a modest 2 percent of the target market in two years. The target population was estimated to be 250,000 physicians. The market potential was .02 times 250,000 or 5,000 customers. In reality, getting 1/2 of 1 percent proved to be an impossible task. After burning

through \$12 million in venture capital over four years, the firm was sold to a competitor for \$1 million.

The third forecasting approach is referred to as causal methods. This approach includes “econometric models. . .to build a model that explains what went on historically and requires identifying the best model to explain and replicate historical demand patterns. For example, if a product is sold mostly to young, affluent singles, an econometric model might be developed to replicate what would happen to demand as they get older, wealthier, and potentially married” (Lapide 2002). This approach can be used successfully for new ventures but the statistical techniques required and the large amounts of data put the technique beyond the reach of most entrepreneurs.

The fourth approach, judgmental methods, involves gathering qualitative and/or quantitative data from stakeholders, such as potential customers, then applying research methods to assess market size or speed of adoption. This approach is the one typically best suited to forecast new products and services but is prone to overstatement of buying intentions by various stakeholders. A forecaster in the pharmaceutical industry stated, “No one has ever been able to adjust for this overstatement. The industry has historically used simple rules of thumb like divide by two or divide by three because it was the only solution available” (Ziment 2004).

This challenge is not unique to the pharmaceutical industry and partially explains why entrepreneurs are less than enthusiastic when someone suggests they use a judgmental method, such as concept testing, as a first step in forecasting sales. In addition, if an entrepreneur does employ this approach and then adjust the results by dividing by two, it may appear as an arbitrary adjustment by external reviewers and result in reduced confidence in the quality of the forecast. The process presented in this article is built on a judgmental forecasting approach but simplifies the analysis and incorporates data warehousing methodology (Inmon 2002) to increase accuracy while reducing complexity. The process also eliminates the “divide by two or three” rules of thumb, which seem so capricious.

Surveying Potential Customers Informally

An entrepreneur was thinking about creating a variation on the in-home personal chef service that had become popular in a small but affluent consumer segment. Her service would be called Your Prep Chef, offering a similar service except the customer would perform the final preparation of the meal. While her concept is similar to the in-home personal prep chef concept, it would likely appeal to a much different audience. Even if it were sufficiently similar, getting accurate data on the existing market would be extremely difficult.

Generally, a budding entrepreneur begins by assuming that everyone will find the concept as compelling as he or

she does. Exuberant optimism is necessary for these ventures to succeed, but at times it blinds the entrepreneur to the realities of the business world. An informal survey at this stage often provides a relatively objective measure and a source of valuable data during the idea development stage.

This particular entrepreneur mentioned the concept to a few coworkers who thought it was a great idea. This pre-screening is an example of a judgmental forecasting approach using an informal concept survey. It is highly encouraged as a first step as long as it is followed by a more formal analysis (assuming the idea passes the informal test). In spite of the favorable response from coworkers, the key question remained, “Will a sufficient number of households purchase the service to make the concept financially viable?” To put it more bluntly and pragmatically, “Should I quit my day job?”

Using an informal judgmental approach tends to be limited to asking family, friends, coworkers, and neighbors for their opinions. Unfortunately, these individuals form a biased sample. They tend to be supportive of the person rather than the concept, which reinforces a “can’t lose” attitude seen in many entrepreneurs. One valid criticism often stated against using a formal concept at this stage of the start-up is exposing the idea to a larger audience that could result in someone “stealing” the idea. While exposing the idea is a risk, it is usually minimal and the benefits more than compensate for the possible loss of future business.

Step 1: Gather the Primary Data with a Formal Survey

The Your Prep Chef (YPC) is detailed to demonstrate the basic elements of the simplified concept survey and data analytic process advocated in this article. Before the “do-it-yourself” meal preparation concept became the “hot” new franchising opportunity (Wilson 2006), a local entrepreneur developed the following idea:

Your Prep Chef will deliver fresh food and recipes for five to seven dinners per week to a customer who would do the actual preparation. A meal could be easily prepared in less than 30 minutes but would be healthy and well balanced. The menus would vary from week to week to add an element of adventure and excitement to the meals. The ingredients would be labeled to correspond to each recipe to simplify the meal preparation process.

The first step in creating an accurate new product forecast is to create a survey and gather responses from randomly selected, impartial individuals. To accomplish this task, the entrepreneur needs to first complete a few simple steps, as outlined below.

- *Define the concept so that it can be clearly communicated in less than 15 seconds.* Many entrepreneurs have trouble describing their concept in clear, simple language. This step is often quite difficult for an entrepreneur since the concept creator can see many variations that inhibit a succinct description. The most important aspect of this step is to test the communication on a number of listeners who are unfamiliar with the concept. An essential part of this step is to observe carefully listener reactions to the description and modify it to quickly communicate the key elements. This description is the starting point of the concept questionnaire.
- *Identify the key features that distinguish the concept from existing or alternate offerings.* How does the product or service meet a customer's need or fill a void in the market? This step includes hypothesizing who (age, gender, education, life style, etc.) will be most inclined to buy.
- *Determine the "product, place, price, promotion" variables that can be "tweaked" to maximize the potential for success.* Items can include size, color, location, price, product quality, etc.
- *Create a one-page questionnaire with the 15-second concept description printed at the top.* Questions that delineate the implementation variables should be placed in the middle of the survey and demographic/target market questions at the end. Construct questions with permissible responses. Avoid open-ended questions to simplify the collection and analysis of the data. Keep the demographic questions to a few relevant items, but always include the respondent's age, gender, and zip code to link the data with data from other sources.
- *Gather around 200 responses.* While telephone interviews are often used, in-person interviews are preferable. The respondent's demeanor and nonverbal cues provide a valuable indication of the person's real feelings about the concept and should be noted.

The actual survey for this analysis can be found in Figure 1.¹

**Table 1. Sample Listing of Survey Data
Table for YPC**

Buy (A)	No. (B)	Fresh (C1)	Ease (C2)	Menus (C3)	Outside (D)	Scratch (E)	Zip (I)
3	2	5	5	3	5	0	40217
3	3	5	3	4	2	0	40206
2	5	5	4	2	4	4	40242
4	0	5	5	1	4	0	40207
4	0	4	4	3	3	1	40065
4	0	5	4	2	2	0	40065
3	2	4	3	2	2	0	40299
5	0	5	3	3	3	0	40299
3	3	5	4	3	7	0	40272
5	0	5	5	3	7	0	40272
4	0	5	3	1	7	0	40218
4	0	5	4	2	2	0	40205
2	2	5	4	3	3	2	40291

Step 2: Create the Primary Data Table and Perform "Top-Level" Analysis

Once the data have been collected, they are entered into a spreadsheet with one row corresponding to a particular respondent and each column representing the individual questions. Microsoft Excel is a good software choice for this procedure because it contains a number of analysis tools that provide a level of sophistication beyond the one presented here for those who are so inclined.

Indexing each questionnaire by placing a sequence number on the form is recommended for control purposes. It is also recommended that the date, time, and location be recorded. These items can be used to cross-reference the data on the sheet to data entered into a computer. Table 1 shows a few of the rows for the YPC concept from the 184 responses gathered. This table, one of two required for the analysis, will be referred to as the survey data table.

Column A is the intent to buy response based on a five-point scale where 1 = definitely buy and 5 = definitely not buy. Columns B through E are shown but are not necessary for the data analytic process. These data were gathered to assist in determining which product variables are most significant for potential customers. Columns F through H contain the demographic data; they are not shown to conserve space.

The first analysis is to average the "buy" column. Calculate the average response by adding all the responses and dividing the total by the number of surveys. For the YPC concept, the overall average is 3.49 or midway between maybe buy and probably not buy. On average, the concept is not well regarded, but averages can be misleading. An old adage goes, "If you put your head in an oven and your feet in a refrigerator, on average you are comfortable."

The first pass of the data can be analyzed a little further, although it is not necessary. The data in Figure 2 are shown to demonstrate that survey respondents have definite opinions of the concept even though quite a few individuals responded "maybe" buy to the first question. To compare the actual responses to a "normal" distribution of responses, the table is sorted in ascending order on the intent to buy question, then the number of respondents for each of the five possible responses are counted. If Excel is used for the analysis, the "countif" function can be used to count the cells with a 1, 2, 3, 4, or 5. For this analysis, the counts are shown graphically in Figure 2 as the vertical bars.

The immediate observation is that the responses are more heavily weighted toward the "no" side than the "yes" side. Those on the fence (3 = maybe) are lower than expected in a normal distribution (the line connecting the dots in Figure 2). This graph shows that many respondents have a definite opinion and it is not very favorable. Thirty-one individuals responded with a 1 or 2 while 88 provided a negative

Your Prep Chef

Survey Seq # _____

Date _____ Time _____ Location _____ Interviewer _____

My name is _____. A local company is thinking about offering a new food service and we would like your opinion. I am not selling anything and the survey will take less than two minutes. The service is called Your Prep Chef. The Your Prep Chef would create menus each week that are interesting, wholesome, and easy to prepare. The service would then prepare all the ingredients, which would be individually measured and wrapped for each recipe. The ingredients for a week's worth of meals would be delivered to the home. The meal is prepared by the customer. The recipes are selected so that each meal can easily be prepared in less than 30 minutes. The service offers the convenience of fast or frozen food with the quality of a home-cooked meal.

A. If this service is priced at \$60 per week, which includes all ingredients for 5 meals for 2 people, delivery and preparation instructions and each additional person costs \$20 per week, how likely are you to buy this service?

1. Definitely buy
2. Probably buy
3. Maybe buy, maybe not
4. Probably not buy
5. Definitely not buy

If person responds with 1–3, ask the following question; otherwise, go to question C.

B. If you purchased this service, how many dinners per evening would you need? _____

C. On a scale of 1 to 5, with 1 being not important and 5 being very important, how would you rate the following items.

_____ Freshness and quality of meats and vegetables

_____ Ease of food preparation

_____ Number of menu alternatives such as low sodium, vegetarian, etc.

D. In the last 7 days, how many dinners did you eat outside the home? _____

If response to D is less than 7, ask question E; otherwise go to F.

E. Of the dinners eaten at home, how many were prepared from scratch? _____

F. Is the head of household or the person primarily responsible for the dinner selection?

_____ under 30 years of age _____ 30–45 _____ over 45

G. Gender of respondent: male female

H. What is your educational level?

1. High school
2. 2 years of college
3. 4 years of college
4. More than 4 years of college

I. Zip code of respondent _____

Figure 1. Consumer-Oriented Concept Survey

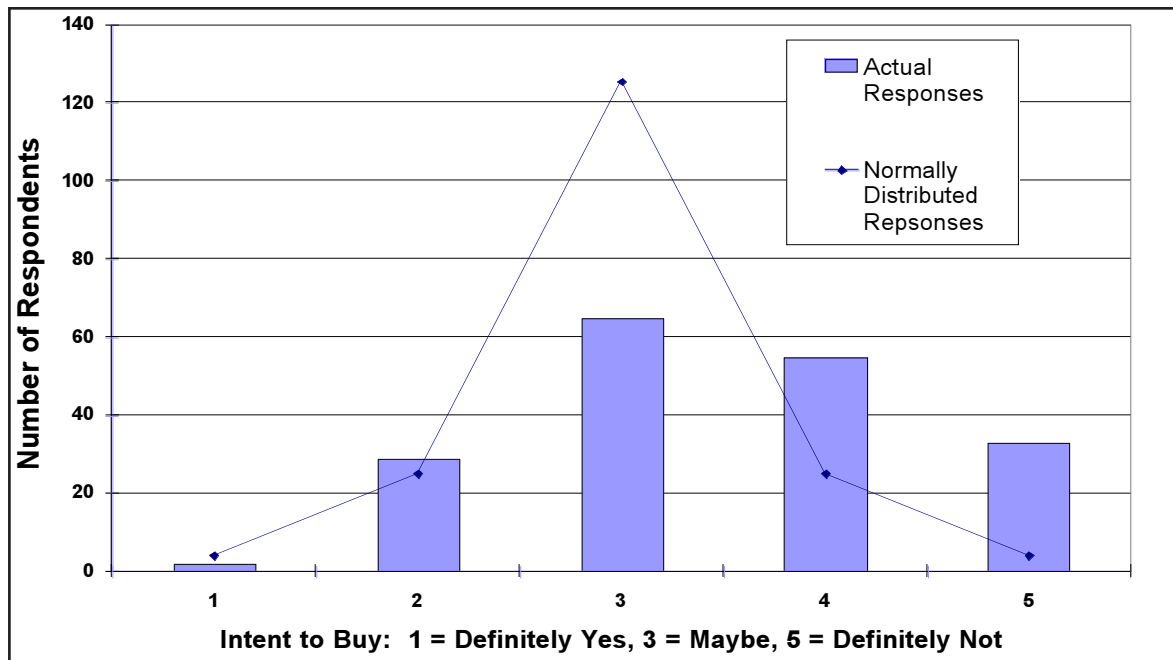


Figure 2. Comparison of Stated Intent to Buy with a Normal Distribution

response. While these results may seem discouraging to the entrepreneur, a little more analysis will yield a much more hopeful outcome.

Step 3: Create the Beginnings of a Data Mart

Of the 184 respondents, 31 responded favorably to the concept. If some common set of attributes can be identified for these 31 individuals, it might be possible to extrapolate or project the survey findings to a larger audience and determine the market size. For instance, say an entrepreneur has a product that he has determined will appeal to males less than 30 years of age who have an advanced degree. The entrepreneur can then focus his efforts on this group by targeting efforts to a very specific group and not waste time trying to market to an audience that is not interested or maybe even antagonistic toward the product.

For the YPC example, the typical gender, age, and education variables were analyzed but do not predict or correlate to purchase intent. In other words, segmenting the data on gender indicates that women are no more likely than men to respond with a definite or probably buy response. The actual analysis is not shown due to space limitations but can easily be performed by sorting on the buy variable then sorting only those responding favorably on gender (or any other variable).

Since none of the more traditional demographic variables lend any useful insights, the original survey data are linked to externally available data (also called secondary data) such as census data. This merging process is typical in creating a data mart, which is a limited or domain-specific data warehouse

(Inmon 2002). For the YPC assessment, “tapestry” groupings developed by ESRI are employed. Derived from census data and grouped by zip codes, “Tapestry segmentation system provides an accurate, detailed description of America’s neighborhoods. U.S. residential areas are divided into 65 segments based on demographic variables such as age, income, home value, occupation, household type, education, and other consumer behavior characteristics” (ESRI 2005).

Three columns are added to the right of the survey data table corresponding to the top three tapestry segments as defined by ESRI. For each zip code, the top three segments are added to the table. Table 2 shows the table after the columns have been added. Columns C through H are hidden to simplify the table. Thus, the original survey data table is merged with secondary data.

Step 4: Data Analytics and Drill Down

A second table, which will be referred to as the tapestry tally table, needs to be created to search for hidden gems in the data. Since its creation appears a little complicated, it will be described in some detail. The goal is to determine if there are pockets of support for the concept that are not obvious from the top-level analysis. In other words are there any tapestry segments that seem to occur more frequently for those responding favorably to the concept? Also, are there some tapestry segments that occur more frequently for respondents who found the concept unattractive?

To answer the above questions, the survey data table is sorted in ascending order on the “buy” variable. Then for all the rows with a 1 or 2 in the buy column, the entrepreneur

tallies or counts the number of times that each tapestry value occurs in any of the three tapestry columns. The tapestry tally table has one row for each of the 65 possible tapestry segments. The first column identifies the tapestry number, the second column is a count of all the times the tapestry number is found in the survey data table where the buy response is 1 or 2. The third column is created by counting the times a particular tapestry number occurs in a row of the survey data table where the buy response is a 4 or 5. When completed, the tapestry tally table appears as shown in Table 3.

The counting process can be performed very quickly and more accurately by using the "frequency" function built into Excel. This function will automatically count the number of occurrences of a particular item, in this case the values from 1-65, in a range of a cell. More information on the frequency function can be found in the help section of Excel.

Once the tapestry tally table is created, the raw counts are converted to percentages by creating three more columns, one for the positive column, one for the negative column, and one for the net percentage. This net percentage column distinguishes this process from most by providing data that eliminate the rule-of-thumb adjustments such as divide by two or three. The goal of this step is to determine those locations, namely zip codes, where the buy sentiment is greater than the not buy sentiment (on a percentage basis).

To determine the percentages for each tapestry, the total number of respondents for the "1 or 2" column and the "4 or 5" column are computed by summing up each column. The percentages are calculated by dividing the count in each cell

by the corresponding total for the column. Once these two columns are created, the last column is calculated by subtracting the "4 or 5" percentage from the "1 or 2" percentage. This net difference is positive in tapestry segments where the sentiment is generally favorable and negative number in tapestry segments where the intent is unfavorable. Table 4 shows the results for all tapestry segments.

A quick scan of Table 4 indicates that tapestry segments 4, 12, 19, 22, and 30 contain the most positive net sentiment while 3, 17, 26, 32, and 33 contain the most negative sentiment. Identifying those zip codes where the sum of the top three tapestry percentages produces the most positive net percentage will provide the first indication of "hidden gems" in the data. This calculation is made by returning to the survey data table as shown in Table 3 (the original survey data must be maintained in ascending order on the buy column if one uses the frequency function to compute the tapestry column data). The net percentage values from the tapestry table that correspond to the tapestry number in the survey table are summed. The resulting table is in Table 5. This data look-up can be automated by using the Vlookup function in Excel.

Table 5 contains only the first 25 rows of the survey data table. A scan through the entire table reveals about 10 zip codes with quite positive responses to the intent to buy question. A further scan of the table indicates that 4 of the 10 zip codes have 10 or more respondents, increasing the validity of the results.

Step 5: Projecting the Data and Identifying a High Potential Target Market

The final step of the process is to take the higher potential

Table 2. Survey Data Table Extended with Tapestry Data

Buy (A)	Zip (I)	Tap 1	Tap 2	Tap 3
2	40014	6	2	19
3	40014	6	2	19
3	40014	6	2	19
3	40023	17	7	2
4	40026	6	2	17
5	40047	26	17	25
3	40059	2	3	0
4	40059	2	3	0
5	40059	2	3	0
4	40065	60	17	13
4	40065	60	17	13
3	40071	17	26	53
4	40071	17	26	53
3	40165	17	26	41
5	40165	17	26	41
5	40203	64	65	62
2	40204	48	22	8
3	40204	48	22	8

Table 3. First 15 Rows of the Tapestry Tally Table

Tapestry Segment	Responses of	
	1 or 2	4 or 5
1	0	2
2	4	10
3	1	12
4	2	1
5	0	0
6	2	7
7	7	23
8	1	2
9	0	2
10	0	0
11	0	0
12	2	1
13	5	10
14	9	23
15	0	0

Table 4. Complete Tapestry Tally Table with Net Intent to Buy

Tapestry Segment	Responses of		Percent of Total		Net Sentiment
	1 or 2	4 or 5	1 or 2	4 or 5	
1	0	2	0.0%	0.8%	-0.8%
2	4	10	4.3%	3.8%	0.5%
3	1	12	1.1%	4.5%	-3.5%
4	2	1	2.2%	0.4%	1.8%
5	0	0	0.0%	0.0%	0.0%
6	2	7	2.2%	2.7%	-0.5%
7	7	23	7.5%	8.7%	-1.2%
8	1	2	1.1%	0.8%	0.3%
9	0	2	0.0%	0.8%	-0.8%
10	0	0	0.0%	0.0%	0.0%
11	0	0	0.0%	0.0%	0.0%
12	2	1	2.2%	0.4%	1.8%
13	5	10	5.4%	3.8%	1.6%
14	9	23	9.7%	8.7%	1.0%
15	0	0	0.0%	0.0%	0.0%
16	1	0	1.1%	0.0%	1.1%
17	0	6	0.0%	2.3%	-2.3%
18	8	20	8.6%	7.6%	1.0%
19	6	8	6.5%	3.0%	3.4%
20	0	0	0.0%	0.0%	0.0%
21	0	0	0.0%	0.0%	0.0%
22	11	26	11.8%	9.8%	2.0%
23	0	0	0.0%	0.0%	0.0%
24	0	0	0.0%	0.0%	0.0%
25	0	1	0.0%	0.4%	-0.4%
26	0	6	0.0%	2.3%	-2.3%
27	0	0	0.0%	0.0%	0.0%
28	0	0	0.0%	0.0%	0.0%
29	8	23	8.6%	8.7%	-0.1%
30	9	21	9.7%	8.0%	1.7%
31	0	0	0.0%	0.0%	0.0%
32	4	16	4.3%	6.1%	-1.8%
33	1	7	1.1%	2.7%	-1.6%

Tapestry Segment	Responses of		Percent of Total		Net Sentiment
	1 or 2	4 or 5	1 or 2	4 or 5	
34	0	0	0.0%	0.0%	0.0%
35	0	0	0.0%	0.0%	0.0%
36	6	14	6.5%	5.3%	1.1%
37	0	0	0.0%	0.0%	0.0%
38	0	0	0.0%	0.0%	0.0%
39	0	0	0.0%	0.0%	0.0%
40	0	0	0.0%	0.0%	0.0%
41	0	3	0.0%	1.1%	-1.1%
42	0	0	0.0%	0.0%	0.0%
43	0	0	0.0%	0.0%	0.0%
44	0	0	0.0%	0.0%	0.0%
45	0	0	0.0%	0.0%	0.0%
46	0	0	0.0%	0.0%	0.0%
47	0	0	0.0%	0.0%	0.0%
48	4	8	4.3%	3.0%	1.3%
49	0	0	0.0%	0.0%	0.0%
50	0	0	0.0%	0.0%	0.0%
51	0	0	0.0%	0.0%	0.0%
52	0	2	0.0%	0.8%	-0.8%
53	0	1	0.0%	0.4%	-0.4%
54	0	0	0.0%	0.0%	0.0%
55	0	0	0.0%	0.0%	0.0%
56	0	0	0.0%	0.0%	0.0%
57	2	4	2.2%	1.5%	0.6%
58	0	0	0.0%	0.0%	0.0%
59	0	0	0.0%	0.0%	0.0%
60	0	2	0.0%	0.8%	-0.8%
61	0	0	0.0%	0.0%	0.0%
62	0	1	0.0%	0.4%	-0.4%
63	0	0	0.0%	0.0%	0.0%
64	0	1	0.0%	0.4%	-0.4%
65	0	1	0.0%	0.4%	-0.4%
Totals	93	264			

zip codes, obtain census household data, and compute the sales potential for these areas. Table 6 contains the results of this final tabulation. The maximum number of households (the last column on the right) is computed by multiplying the Net % Buy for Zip Code column times the Number of Households (HH) in the Zip Code column. The last column is an estimate of how many households would buy if the net percent buy figure is projected for the entire zip code.

At this point the entrepreneur can begin to draw some conclusions about the idea's potential. By drilling down into the data, pockets of opportunity that were not visible in the top-level analysis are identified. The first three zip codes appear to offer a very good opportunity for the entrepreneur. The fourth zip code, 40223, has less potential due to the lower number of households that could be expected to try the service. In all four cases, selecting the "best" location will require additional information such as proximity to available commercial space, ease of delivery in each area,

discretionary spending, and ability to target market in the location.

For discussion purposes, we will assume that the second zip code, 40291 is selected as the best location and that the venture strategy would be to completely saturate this market in one year. In other words, the entrepreneur would like to grow the business in this first location to the 385 households in the first year.

Table 7 contains a revenue projection assuming 25 households can be secured in the first month each paying \$75 per week for the service with an average of four weeks in a month. The total revenue for the first year is \$498,000. While there are a number of "ifs" for the concepts such as "if the expenses can be managed," the concept appears to have significant potential. Reaching 100 percent of potential is extremely difficult, so the above represents a high-end estimate. One could use Pareto's law and assume 80 percent of potential with 20 percent effort as a good top-end estimate.

Table 5. Final Survey Data Table Excerpt with Net Sum of Top Three Tapestry Segments Net Buy Percentage

Buy (A)	Zip (I)	Tap 1 ID	Tap 2 ID	Tap 3 ID	Net % for Zip code
1	40220	13	36	29	2.6%
1	40241	12	4	2	4.1%
2	40299	18	33	7	-1.7%
2	40216	29	32	57	-1.2%
2	40216	29	32	57	-1.2%
2	40213	32	29	36	-0.7%
2	40243	7	6	18	-0.7%
2	40214	48	32	29	-0.6%
2	40207	22	14	3	-0.5%
2	40223	2	30	7	1.1%
2	40220	13	36	29	2.6%
2	40220	13	36	29	2.6%
2	40220	13	36	29	2.6%
2	40291	18	19	7	3.3%
2	40291	18	19	7	3.3%
2	40291	18	19	7	3.3%
2	40291	18	19	7	3.3%
2	40014	6	2	19	3.4%
2	40204	48	22	8	3.6%
2	40242	16	13	18	3.7%
2	40241	12	4	2	4.1%
2	40206	48	22	36	4.4%
2	40205	22	14	30	4.7%
2	40205	22	14	30	4.7%
2	40205	22	14	30	4.7%

Table 6. Top Four Market Opportunities for YPC with Household Census Data

Net % Buy	Zip Code	No. of Surveys	Number of HHs	Residents per HH	Income per HH	Max No. of HHs
4.7%	40205	28	10699	2.2	\$ 63,912	503
3.3%	40291	10	11652	2.6	\$ 66,198	385
2.6%	40220	17	14553	2.3	\$ 57,973	378
1.1%	40223	11	9099	2.5	\$ 75,688	100

Table 7. Monthly Revenue Projections for the 40291 Zip Code

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
# HH/Week	25	32	41	53	68	87	111	142	183	234	300	385
Price/Wk/HH	\$ 75	\$ 75	\$ 75	\$ 75	\$ 75	\$ 75	\$ 75	\$ 75	\$ 75	\$ 75	\$ 75	\$ 75
Sales/Month	\$7,500	\$9,617	\$12,330	\$15,810	\$20,272	\$25,992	\$33,327	\$42,732	\$54,792	\$70,254	\$90,080	\$115,500

Conclusions

The concept survey and the associated secondary data (census data in the above example) contain very useful marketing information that can be used to fine-tune not only the location (place), but also the marketing message (promotion) and the price. If the survey is properly constructed, the fourth P of marketing, namely product, can be perfected by using survey respondents' comments on product characteristics. As can be seen in the sample survey in Figure 1, question C contains three product related characteristics that can assist the entrepreneur in refining the product and the marketing message to better reach and meet the needs of the target audience.

The analysis of the product-related data is beyond the scope of this research, but it is important to include ques-

tions such as this in the early stage of concept analysis. It takes very little time to obtain responses to these types of questions and they can be invaluable in refining the concept, thereby increasing the probability of a successful venture.

The use of customer surveys to refine a concept or determine market potential can be accomplished in an economical manner without abandoning accuracy or precision. Utilizing some of the best practices of data warehousing allows entrepreneurs to develop a sales forecast that is easy to create and to communicate to others while being reliable and accurate. The approach detailed above will demonstrate to potential stakeholders that the entrepreneur has fully thought out the concept and has created a plan built on a solid foundation of accurate information.

Endnote

1. The actual data table is available on request from the first author or at business.bellarmine.edu/ypc.htm.

References

- ESRI. 2005. *Tapestry segments by zip code*. Retrieved December 15, 2005. <http://www.esribis.com/reports/ziplookup.html#about>.
- Inmon, W. H. 2002. *Building the data warehouse*, 3rd ed. New York: John Wiley and Sons.
- Lapide, Larry. 2002. New development in business forecasting. *Journal of Business Forecasting Methods and Systems* 21:1-5.
- Wilson, S. 2006. In the Know. *Entrepreneur* (January): 88-89.
- Ziment, Howard. 2004. Pharmaceutical marketing; new research technique offers improved pharmaceutical product forecasting. *Health & Medicine Week* (May 31): 969.



About the Authors



MICHAEL D. MATTEI (mmattei@bellarmine.edu) is the executive director of the Bellarmine Center for Entrepreneurship and professor of management in the Rubel School of Business at Bellarmine University. Dr. Mattei received his Ph.D. from the University of Louisville. He gained more than 15 years of experience at Cincinnati Milacron, General Electric, and Brown & Williamson where he served as manager of domestic and international marketing systems. Prior to beginning a career in teaching, he was VP of IT for IMX, Inc., a medical technology start-up, funded with \$12 million from the Welsh, Carson VC fund.



STEPHEN HELLEBUSCH (steve@hellrc.com) is president of Hellebusch Research & Consulting, Inc., a 14-year-old company specializing in the pharmaceutical area. He has extensive expertise in the prescription-to-nonprescription switch area. Dr. Hellebusch began his career at Burke Marketing Research, Inc., and went on to Mobil Corporation and Merrell Dow Pharmaceuticals, where he learned the client side of the marketing research business. He received a Ph.D. from the University of Notre Dame.