Columbia Business School
AT THE VERY CENTER OF BUSINESS™

Columbia CaseWorks

ID#070201

PUBLISHED ON
NOVEMBER 22, 2016

Personal Training at the New York Health Club

BY COSTIS MAGLARAS*

Introduction

Tom, a senior manager at the New York Health Club (NYHC), looked over the pile of his personal trainers' time cards stacked on his desk. Personal training comprised a significant revenue stream for the club—one that was becoming increasingly vital to the organization's long-term profitability. Yet Tom was concerned that NYHC was doing a poor job of managing this service. In particular, he felt that the way that NYHC priced personal training was undisciplined and created the potential for fraud. He wondered how much this was costing the club and what could be done to improve the situation.

Industry Overview

At the beginning of the 21st century, the health club industry in the United States was large but fragmented. At the end of 2004 there were 26,046 US health clubs, with a total of 39.4 million members. The median age of those members was 41. Approximately 51% were women 51% and 49% were men. They had an average household income of $61,000.

In 2004 personal training was the strongest growth segment within the fitness industry. Ninety-four percent of all US gyms provided personal training for a range of services including post-rehabilitation training, sports conditioning, weight management and counseling, and prenatal fitness.

Personal Training Program

Trainers at NYHC were responsible for signing up their own clients, who negotiated with their trainer to arrive at a price for the service at one of four standard rates: $50, $60, $80, or $100 per hour. The rate a client paid was based solely on the trainer's ability to negotiate as high a

Author affiliation
David and Lyn Silfen Professor of Business, Columbia Business School

Acknowledgements
James Farrell '06, Alison Jatlow '06, Chris Meron '06, Jennifer Norwick '06, and Erica Pergament '06 provided development support for this case. Professor Garrett Van Ryzin provided development support for the exercises in modeling customer choice.

price as possible; it did not reflect the type of service provided. As a result, prices varied across clients and across trainers.

Revenues from personal training were split 50/50 between the gym and the trainer. The average revenue per session in 2005 was $68.10 ($34.05 each for the gym and the trainer). NYHC had almost unlimited access to qualified personal trainers.

While Tom knew that most of his trainers were quite savvy about negotiating the highest price possible with each client, he had long been concerned about the loose pricing structure at NYHC. He heard from staff that new members often found the process discouraging and intimidating. Moreover, he feared that there was a high potential for fraud because trainers could sign up clients at a low price and then pocket additional money on the side. He wondered if a more disciplined pricing policy could help streamline the selling process and increase the gym's revenues. The simplest idea would be to apply a flat rate for all sessions, as most of NYHC's competitors did.

Another alternative would be to use a system that exploited the variation in demand for personal training services throughout the day. Some times were much more popular than others. Specifically, Tom's data that showed there was a peak in the number of people being trained in the early morning (6 a.m.–9 a.m.) and evening (5 p.m.–9 p.m.) periods (see Exhibit 1). Could prices be lowered during off-peak hours in a way that would increase the demand during these hours and improve overall revenues for the gym? Could this be achieved without impeding the gym's ability to earn high revenues during peak hours? Tom decided that to answer these questions he needed to know more about the personal training preferences of his members. To that end, he created a survey to elicit this information.

Tom's Survey Data

NYHC had 10,000 members. A total of 1,000 of them responded to the survey; of these, 292 (29.2%) were students, 469 (46.9%) had fixed work schedules, and 239 (23.9%) had flexible work schedules. Tom felt that it was reasonable to assume that on average all the members of each group visited the gym the same number of times per week. Respondents' answers to question 1 of the survey, which can be interpreted as measuring the willingness-to-pay for personal training sessions at different times during the day, can be found in NYHCSurvey.xlsx, which is available in a separate file.

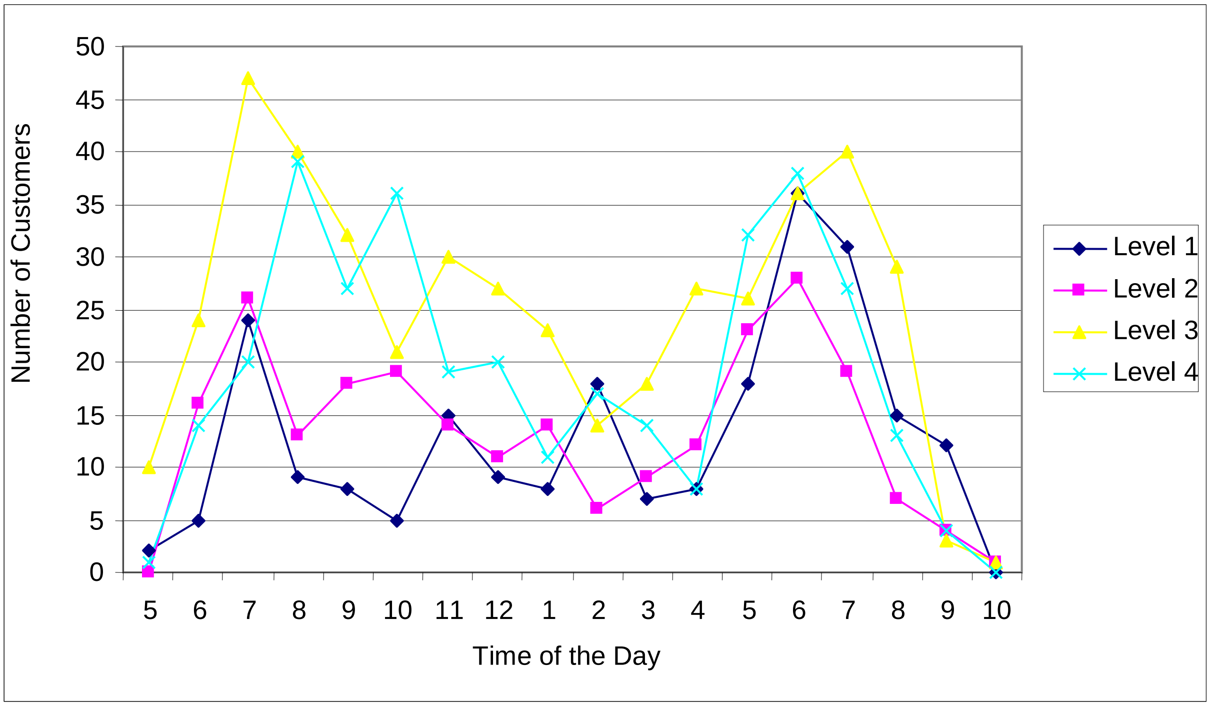
Exhibits

Exhibit 1

Training Session Data

TIMING AND PRICE OF CUSTOMER SESSIONS

(Level 1= $50/hour; Level 2= $60/hour; Level 3= $80/hour; Level 4=$100/hour):



Questions: Part A

1. Suppose that NYHC offered a flat rate for personal training for all members and all time periods. What price would you recommend?

2. Suppose that the gym's strategy is to offer personal training sessions at two price levels: all times other than during the 5 p.m.–9 p.m. slot will be priced at $65 per hour, while sessions between 5 p.m. and 9 p.m. will be priced at some other level, which is to be determined. Under that assumption, construct a linear demand model for the aggregate potential demand from all member segments for personal training sessions between 5 p.m. and 9 p.m.

   By aggregate, we mean the demand from all three member segments, i.e., students, members with fixed schedules, and members with flexible schedules. Members decide which product to buy by first calculating the net utility of a personal training session (i.e., their willingness-to-pay minus the actual cost of the session) at various time slots. They then select the option that maximizes net utility; if all options result in negative utility for a particular member, then that member does not buy any sessions at all. By fixing the price at $65 per hour for all time periods except 5 p.m.–9 p.m., you can compute the net utilities for all these products for each member who participated in the survey. You can then vary the price for the 5 p.m.–9 p.m. slot and, for each possible choice, count the number—or, even better, calculate the fraction—of members who would choose a session in the 5 p.m.–9 p.m. period. You should then try to approximate this demand function by a linear relation of the form d(p) = D – b*p, where d(p) is the fraction of members who purchase a session during the 5 p.m.–9 p.m. slot at price $p, D is the maximum potential demand when $p = $0, and b is the price sensitivity parameter that explains how demand drops as price increases. The goal is to find D and b.

   In addition, construct an exponential demand model for the aggregate potential demand for personal training sessions between 5 p.m. and 9 p.m. from all member segments under the assumption that the prices for all other products is $65 per hour. That is, try to fit a demand function of the form D(p) = D*exp(-b*p), where exp(x) is the exponential function.

3. The demand for the sessions between 5 p.m. and 9 p.m. depends on the price of sessions during all other time periods. Denote by p0 the price charged for all periods other than 5 p.m.–9 p.m., and denote by p59 the price charged from 5p.m. to 9 p.m. Consider a grid of possible pairs of prices (p0, p59) between $35 and $125, and for each such combination of prices compute the resulting demand for the 5 p.m.–9 p.m. time slot. Using this demand information, estimate a linear model for the demand for the 5 p.m.–9 p.m. period, of the form (estimate D0(p0,p59):

   D59(p0, p59) = D59 – b1*p59 + b2*p0.

   And similarly estimate a model for the demand on all other time slots as

   D0(p0,p59) = D0 – b3*p0 + b4*p59

You should estimate the model by running a linear regression between the observed demands D0(p0, p59) and D59(p59,p0) against the two prices (p59, p0); this should be two separate regressions.

4. Formulate a revenue optimization problem where you use the two demand models estimated in question 3 and find the solution to the following two pricing problems.
    a. Fix the price p0 = $65, and pick the price p59 so as to optimize the total revenue extracted from all personal training sessions across all time slots.
    b. Optimize over both p0 and p59 to maximize the total revenue extracted by NYHC, incorporating the constraint that p0 <= p59.

Questions: Part B

Tom knows that he can efficiently identify the students (by requiring a valid student ID), and he is interested in seeing if it would be worthwhile for NYHC to offer a separate menu of prices for personal training for students.

1. To better understand the student preferences regarding personal training:

    a. Compute the average willingness-to-pay for student members for each time slot.

    b. Plot willingness-to-pay histograms for each of these time slots.

    c. Do you think it would be sufficiently accurate for purposes of student demand estimation and revenue optimization to restrict attention to a model that summarized the student preferences by their average willingness-to-pay for each time slot? Specifically, assume that you are given a vector of prices and are asked to compute the student demand for personal training sessions at different periods. Do you think it is reasonable to assume that all students have the same willingness-to-pay (i.e., the average computed above in 1a) or is the variability in their willingness-to-pay an important element in this calculation? Why or why not?

The following questions focus on modeling the choice behavior of the NYHC members with respect to personal training sessions using a multinomial logit model (MNL). See Appendix for some background on the multinomial logit model. Willingness-to-pay data can be found in NYHCSurvey.xlsx.

2. Fit a MNL model to the student and non-student segments of NYHC members based on the reported WTP survey data. The non-student segment is all residual members with fixed and flexible schedules. Clearly identify the model parameters.

3. Suppose NYHC decided to offer two flat rates, one for student members, and the other for all other members of the gym. The flat rates would apply to all time slots. For example, the student flat rate would price equally personal training sessions for students across all 6 time slots. Combine the MNL demand models for the student and the non-student segments to get an aggregate demand model. You may assume that the fraction of student members is similar to that of the survey, that is, 29.2% of the members are students. Use this aggregate demand model to determine the prices that would maximize the gym's daily revenue from personal training sessions. The student price should be less than or equal to the general member price.

4. While there is unlimited supply of personal trainers that could be hired, if needed to fulfill demand, Tom realizes that the gym starts to feel congested if there are more than 20 private sessions on the floor at each hour. Consider the problem in question 3 and assume that when a member decides to purchase a private training session on a particular time slot, this really implies that she or he will purchase 1 session per week

at that time slot, and the particular day-of-the-week in which this session will actually be held is equally likely to be a Monday, Tuesday, …, Sunday; and the specific hour on which the session will occur is also uniformly distributed in the respective time slot. For example, if the aggregate demand of the previous question resulted in 400 personal training sessions across the week for the 12pm-2pm time slot, then this would correspond to 57 sessions per day over that time slot, and 28.5 sessions per hour of that time slot.

a. What is the per hour utilization of the gym by personal training sessions under the pricing proposal you determined in question 3? What is the breakdown by student members and non-student members? (Suppose NYHC closes at 11PM every day.)
b. Assume that Tom still wants to pick a flat price for students and another flat price for non-students (i.e., the same price for all time slots per segment), and wants to ensure that the expected demand in each hour is less than or equal to 20.
   i. What are the pair of optimal prices? Again, the student member price should be less than or equal to the general member one.
   ii. What is the per hour utilization of the gym from personal training sessions? What is the breakdown of this utilization per segment?

5. Consider the same setup as in question 4. For the student members, Tom wants to charge a flat price for personal training sessions across any time slot. For general members, he is willing to designate some time slots as "peak" and the rest as "non-peak" and charge one price for all peak time slots and another for all non-peak time slots. Would that help? If so, identify a set of peak time slots (you do not have to exhaustively try all possible combinations) trying to justify the rationale behind your choice. What are the resulting optimal pricing? What is the utilization per hour, and its breakdown by segment?

Appendix

Background on Multinomial Logit Model

Suppose there are many products each consumer can choose from. In the NYHC case, gym members can choose between 6 products, one for each time slot. For each consumer k, his/her willingness-to-pay (WTP) for product i is $u_i + \epsilon_i^k, i = 1,2, \dots, 6$. Specifically, u_i is equal to the average WTP and is assumed to be common across all consumers of the same segment (e.g., all student members). And, ϵ_i^k is a random, consumer-specific WTP perturbation around the average value, which differentiates consumers among themselves. ϵ_i^k follows a Gumbel distribution with zero mean and shape parameter μ (the shape parameter is proportional to the standard deviation of these random WTP perturbation terms, and specifically the $\mathrm{Stdev}(\epsilon_i^k) = \pi\mu/\sqrt{6}$). For each consumer k, the ϵ_i^k are independent, identically distributed across all products i, $i = 1,2, \dots, 6$, and they have the same μ (i.e., standard deviation). [Note: this "implies" that in practice we cannot have very dissimilar products in the choice set, e.g., some that are worth \$100 ± random variation, and some that are worth \$1000 ± random variation, because we are assuming that the random variations are of "similar" magnitude (same standard deviation).] The ϵ_i^k are also independent across consumers k.

For the model above, the probability that a consumer k chooses product i is

$$P(\text{choose i}) = P\left(u_i + \epsilon_i^k - p_i \geq u_j + \epsilon_j^k - p_j, \text{ for all } j \neq i, \text{ and } u_i + \epsilon_i^k - p_i \geq 0\right) = \frac{v_i}{1 + \sum_{j=1}^{6} v_j},$$

where

$$v_i = e^{\frac{u_i - p_i}{\mu}}, \quad i = 1, \dots, 6,$$

v_i can be interpreted as the "attractiveness" of product i, and the probability of choosing i among an option set is proportional to the attractiveness of i over the attractiveness of the entire option set $1 + \sum_{j=1}^{6} v_j$; where $1 = e^0$ = attractiveness of the "no purchase" option that is worth \$0 of utility.

The demand for product i from a market of size N, say N=10,000 in this case (if all gym members belonged to the same segment), is

$$\text{demand(product i)=} \quad d_i(p) = N \cdot \frac{v_i}{1 + \sum_{j=1}^{6} v_j} = N \cdot \frac{e^{\frac{u_i - p_i}{\mu}}}{1 + \sum_{j=1}^{6} e^{\frac{u_j - p_j}{\mu}}}$$

Finally, one way to extract the MNL model parameters from a survey that generated a sample of WTP data for the different offered products is to set the u_i = sample average of the reported WTPs for product i, and set the common shape parameter as follows: First compute the variance of the WTP reported for each product i, which we will denote by V_i. Then compute the average across these variance terms $V = (V_1 + \ldots + V_I)/I$. And then set:

$$\mu = \frac{\sqrt{6V}}{\pi}$$

[More generally, we may also consider how is the average WTP u_i explained by the features of a product, e.g., the size of the screen of a phone, its memory, its camera characteristics, etc.

$u_i = \sum_{f:features} a_i^f x_i^f$, where x_i^f = value of that feature f for product i, and a_i^f = "part worth" or value of that feature. In that case,

$$d_i(p) = N \cdot \frac{e^{[(\sum_f (a_i^f x_i^f) - p_i)/\mu]}}{1 + \sum_{j=1}^{6} e^{[(\sum_f (a_j^f x_j^f) - p_j)/\mu]}}.$$

The x_i^f are the known product features, and the a_i^f are the model parameters that we would try to estimate to "explain" the reported WTP numbers or the observed choices of a sample of consumers. In NYHC, we do not have the detailed features of each gym session slot.]