PREDICTING CUSTOMER CHURN AT QWE INC.

QWE Inc.[1] helped small- and medium-size businesses manage their online presence through a subscription service. As with many successful dot-com start-ups, QWE experienced fast growth initially but, as the company and its business model matured, management realized the need for deeper analytical insight into some key business processes, one of which was customer retention.

QWE customer contracts ran either month to month or for a fixed period of 6 or 12 months. Thus, depending on the type of contract, a customer was free to leave either at the end of each month or at the end of Month 6 or Month 12 of its relationship with QWE.

Up until 2012, QWE took a reactive approach to *churn*, or attrition: if a customer called with a request to cancel his or her contract, the QWE customer service representative would try to convince the customer to extend the contract, most often by offering free services or discounts on existing services. In February of that year, Richard Wall, VP of customer services at QWE, wondered if his team could develop a more proactive approach.

Wall thought that if QWE could estimate the probability that a given customer would leave in the near future and identify the drivers that contributed most to that customer's decision, then his team would be able to reach out to the customer, enhance his or her experience with QWE services, and divert churn without giving up costly discounts. To help his team prioritize its efforts, Wall decided that the goal should be to use current characteristics to determine probability of churn within two months.

He asked one of his associates, V. J. Aggrawal, a recent graduate of a respected business school, to take on this project. Wall wanted Aggrawal to generate a list of the 100 customers who were most likely to leave and, if possible, the three factors contributing most to that likelihood.

---

[1] The company name and the names of individuals are disguised to protect confidentiality. The data are real; data are adjusted by an unspecified constant so that all relationships are preserved.

This field-based case was prepared by Anton Ovchinnikov, Assistant Professor of Business Administration. It was written as a basis for class discussion rather than to illustrate effective or ineffective handling of an administrative situation. Copyright © 2013 by the University of Virginia Darden School Foundation, Charlottesville, VA. All rights reserved. *To order copies, send an e-mail to* sales@dardenbusinesspublishing.com. *No part of this publication may be reproduced, stored in a retrieval system, used in a spreadsheet, or transmitted in any form or by any means—electronic, mechanical, photocopying, recording, or otherwise—without the permission of the Darden School Foundation.*

Data Collection and Relevant Customer Characteristics

Honored but intimidated by the task, Aggrawal first researched what data was available. He logged onto QWE's SalesForce.com database, rolled back two months to December 1, 2011, and obtained a sample of 6,000 of QWE's customers as of that date. He found the number of customer characteristics overwhelming—there were more than 230 data fields for each customer. So he approached Wall, hoping to tap into his intuition about which characteristics were most important.

"Customer age, CHI [Customer Happiness Index], and service and usage patterns. I'd probably start with those," said Wall. "Regarding age, I'd say that if a customer has been with us for more than 14 months, he or she knows how to use our services, is using them, gets value out it, and should be less likely to leave. Those who have been with us for less than 6 months are still perhaps only learning about the services, so I am not sure how it will go, but those between 6 and 14 months are probably the riskiest group."

"Makes sense," nodded Aggrawal.

"We do measure customer happiness," continued Wall. "CHI combines multiple characteristics related to customers' experiences with our services. I doubt that those with high CHI scores leave much, but those who are unhappy might leave, and so might those for whom CHI scores dropped recently."

"Yes, yes. I remember we discussed that in our team meeting last week," responded Aggrawal.

"Now for service," said Wall, pausing at length before proceeding. "I am least sure about this one, but something tells me that any customer using our system may have some issues at times. When requests to resolve those issues are routed to the tech people, these requests are assigned a priority level depending on how serious the issue is. So when a customer has many service requests, or if one has a high priority—meaning the issue is serious—the customer may indeed have some problems with our services and may just drop because of that. On the flip side, if a customer is actively using the system, logs in a lot, writes blogs, and sees that others pay attention to those blogs, that customer likely sees value in our services, so why would he or she leave?"

"Thank you so much, Richard! This is tremendously helpful," agreed Aggrawal, who then moved on to coding and filtering data.

Aggrawal used *Month 0* to denote the current moment in time. *Month 1* was one month before, *Month 2* was two months before, and so on. Deciding to start with something simple, he pulled only the data for November and December 2011, and as Wall suggested, recorded customer longevity in months, CHI score, number of support cases, average support priority, and usage information: logins, blogs, views, and days since last login.

Following Wall's advice, Aggrawal pulled data on both the value of a characteristic as of December 1, 2011, and its change from November to December. **Exhibit 1** presents a snapshot of the first 10 customers in his sample.[2]

Finally, for each customer in the sample, Aggrawal added a column indicating whether the customer actually left in the two months following December 1. Ultimately, he was hoping to relate the information about a customer, known as of December 2011, to the likelihood of that customer leaving before February 2012.

---

[2] For the entire sample, see the supplementary spreadsheet provided with this case, UVA-QA-0806X.

Exhibit 1

PREDICTING CUSTOMER CHURN AT QWE INC.

Data Sample

ID	Customer Age (in months)	Churn (1 = Yes, 0 = No)	CHI Score Month 0	CHI Score 0-1	Support Cases Month 0	Support Cases 0-1	SP Month 0	SP 0-1	Logins 0-1	Blog Articles 0-1	Views 0-1	Days Since Last Login 0-1
1	67.0	-	-	-	-	-	-	-	-	-	-	31.0
2	67.0	-	62.0	4.0	-	-	-	-	-	-	(16.0)	31.0
3	55.0	-	-	-	-	-	-	-	-	-	-	31.0
4	63.0	-	231.0	1.0	1.0	(1.0)	3.0	-	167.0	(8.0)	21,996.0	-
5	57.0	-	43.0	(1.0)	-	-	-	-	-	-	9.0	31.0
6	58.0	-	138.0	(10.0)	-	-	-	-	43.0	-	(33.0)	-
7	57.0	-	180.0	(5.0)	1.0	1.0	3.0	3.0	13.0	(1.0)	907.0	-
8	46.0	-	116.0	(11.0)	-	-	-	-	-	-	38.0	6.0
9	56.0	-	78.0	(7.0)	1.0	(2.0)	3.0	-	(9.0)	1.0	-	7.0
10	56.0	-	78.0	(37.0)	-	-	-	-	(7.0)	-	30.0	14.0

Data source: Disguised company data.